ESSAYS IN ENERGY AND ENVIRONMENTAL ECONOMICS

A Dissertation

by

JONATHAN BARRETT SCOTT

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Steve Puller
Committee Members,	Mark Hoekstra
	Fernando Luco
	Reid Stevens
Head of Department,	Timothy Gronberg

August 2018

Major Subject: Economics

ABSTRACT

This dissertation includes three essays related to energy and environmental economics, with an emphasis on policy. Topics discussed pertain to policies aimed at internalizing externalities, or encouraging adaptation to climate change. I discuss these topics in the context of the transportation sector, natural disaster insurance, and infrastructure policy.

I analyze consumer behavior in a vehicle arms race—where preferences for safety lead to strategic purchases of large vehicles—as a potential contributing factor to the rise in vehicle sizes in the U.S. First, I test for arms race purchasing behavior by exploiting quasi-random next-door neighbor fatalities in fatal car accidents in order to examine how shifts in preferences for safety impact demand for heavy vehicles. I find that households neighboring an individual who dies in an accident respond by purchasing significantly larger and safer vehicles than households neighboring someone who survives an accident. Second, I explore how gasoline taxes counteract arms race behavior by estimating a discrete choice model of vehicle purchases, allowing for consumer preferences on *relative* weight. My specification allows preferences for vehicle size to vary based on the size of cars in the households' area. Counterfactual simulations illustrate that an arms race can be reversed with a sufficiently high Pigouvian tax, which will depend on the distribution of vehicle weights, specific to the area.

The next study examines the unintended role government policies can play in discouraging climate change adaptation. Despite the large costs of covering flood losses, little is known about whether flood insurance availability affects the decision to live and stay in more flood-prone areas. In this paper, we present evidence that suggests households in flood-prone areas would have otherwise moved to less risky areas absent flood insurance availability. We identify the effect of flood insurance availability on population flows by exploiting the within- and across-county variation in the various programs that the federal government implemented to encourage flood-prone areas to join

the National Flood Insurance Program (NFIP). Results suggest that flood insurance availability caused population to increase by 4 to 5 percent in high flood-risk counties. Furthermore, we find that NFIP causes a 4.4% increase in population per one standard deviation increase in risk.

Finally, I analyze how infrastructure policy may encourage coal plants to convert to clean natural gas energy. The U.S. Environmental Protection Agency's (EPA's) Mercury and Air Toxics Standards (MATS) have imposed significant costs on coal-fired power plants in recent years. To many old, inefficient coal plants, high costs of compliance have prompted a trend toward conversions to natural gas—a much cleaner source of energy. Transportation frictions may inhibit some of these conversions from taking place. In this paper, I analyze the role that the natural gas pipeline infrastructure plays in incentivizing these conversions. Exploiting MATS' increased pressure on coal electric generating units (EGUs) with capacity 25 megawatts or greater within a triple-differences framework, I estimate that a 10% reduction in pipeline distance produces an additional 6.4% increase in likelihood of conversion for an EGU regulated under MATS. Isolating the impact of pipeline costs within a dynamic model of the plant's decision to convert (still fully exploiting MATS), I find that pipeline subsidies alone can produce one-third of the conversions as emissions reduction mandates under MATS, and present value of external benefits of $10.1 billion from reduced emissions. In addition, I estimate a marginal benefit of $2.6 million per-mile of pipeline.

ACKNOWLEDGMENTS

I am grateful to my committee chair, Steve Puller, for his continued guidance and support. If not for his valuable advice and encouragement I would not be where I am today.

I am grateful to my dissertation committee: Mark Hoekstra, Fernando Luco, and Reid Stevens for further research advice and support. Additionally, I thank everyone in the applied microeconomics group at Texas A&M for comments throughout countless seminars.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Steve Puller, Mark Hoekstra, and Fernando Luco of the Department of Economics, and Professor Reid Stevens of the Department of Agricultural Economics.

The analysis in Section 3 was conducted in part by Abigail Peralta of the Department of Economics at Texas A&M.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by assistantships from Texas A&M University, a research fellowship from Strata Policy, and a dissertation research fellowship from Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Externalities are a primary driver of market inefficiencies. These externalities are often produced as a byproduct to consumer, producer, or even government behavior pertaining to market activities. In this dissertation, I study environmental externalities produced by each one of these entities. The emphasis in this research is on policy solutions which would internalize some or all of the external costs produced by these actions. In each of the following chapters, I study and/or offer policy solutions to these environmental externalities, in order to achieve more efficient economic outcomes. This analysis will concentrate on three sectors: transportation, government subsidized flood insurance, and electricity.

In Chapter 2, I examine costly externalities produced by vehicle consumers in the transportation sector. Specifically, I study one potential mechanism which has caused increases in the size of the fleet in the past several decades. This factor—described as a vehicle arms race—occurs when consumers strategically purchase larger vehicles in order to protect themselves from other cars on the road. These strategic responses are, therefore, a function of consumer preferences for safety. The literature on the vehicle arms race is quite sparse, and there is little to no causal evidence of these "arms race"-*style* preferences. In this chapter, I isolate shocks to preferences for safety in the form of household responses to the death of a *next-door* neighbor in a vehicle accident. Using a unique household-level dataset with exact addresses, I exploit the quasi-random nature of a neighbor's fatality using within household variation in purchases. My estimates suggest that households respond to this shock to preference for safety by increasing their vehicle sizes significantly over time. Since in a vehicle arms race, these preferences for safety are presumably linked to the size of other vehicles on the road—generating feedback loops in fleet size—I directly examine this relationship within a discrete choice model of vehicle purchases. Specifying preferences for vehicle size as a function of consumers' local distribution of vehicle weights in their area allows me to estimate the impact of several counterfactual policy interventions which would serve to internalize this feedback loop. I examine a Pigovian gas tax and illustrate that with a sufficiently high tax an arms race can be reversed. That is, raising gasoline prices via a tax induces households to purchase more fuel efficient (i.e., smaller) vehicles, which reduces the size of vehicles on the road, and thus induces the next round of purchasers to buy smaller cars.

In Chapter 3, co-author Abigail Peralta and I look at federally subsidized flood insurance from the National Flood Insurance Program (NFIP), and analyze the impact of the poor incentives it produces. In the face of rising sea levels and increased risk of flood, adaptation may be a natural response for households. For example, households might adapt by migrating out of the areas that are more susceptible to flood. However, when faced with highly subsidized flood insurance, being compensated for potential losses could shift households on the margin away from moving out of a flood-zone. We estimate a treatment effect of receiving NFIP on the population of the treated, risk-prone areas of the U.S. In this chapter, we exploit within- and across-county variation in various, plausibly-exogenous, programs that the federal government implemented to encourage flood-prone areas to join the NFIP. We demonstrate that the counterfactual response—in absence of flood insurance take-up—would be a decline in the population of these high flood-risk areas. Specifically, we estimate that NFIP is responsible for a 4 to 5 percent increase in population in high flood-risk counties. In addition, we study the differential effects of subsidized flood insurance availability based on local area's relative flood risk. We find that NFIP causes a 4.4% increase in population per on standard deviation in risk. This finding suggests that NFIP contributed to a 6.6% increase in costs attributed to hurricane Katrina, and up to a 15% increase in costs associated with hurricane Harvey. The moral hazard produced by the NFIP, and uncovered in this chapter, provides strong evidence that premiums are much too low in the country's most flood-prone areas.

Chapter 4 of this dissertation studies the costly externalities produced by the country's coal-fired power plants, and looks to infrastructure as a possible factor that could reduce these external costs. This chapter examines the relationship between the natural gas pipeline infrastructure and natural gas usage. Newly imposed regulations, such as the Environmental Protection Agency's (EPA's) Mercury and Toxics Standards (MATS), have placed significant costs on power plants generating electricity from coal or oil. The primary option for the affected plants is to invest in costly equipment to achieve compliance. However, a recent trend is the conversion from coal or oil to natural gas fired-burners. These transitions to natural gas usage have enormous environmental benefits, as burned natural gas emits about half the carbon dioxide of coal. In this chapter, I exploit as-good-as-random variation in MATS treatment—imposed on coal and oil electric generating units (EGUs) of 25 megawatts or greater—in order to estimate the influence that proximity to pipeline has on the plants' decision to convert an EGU subject to MATS. The estimates presented in this chapter demonstrate that MATS creates significantly larger incentives for coal plants to convert to natural gas when the plant is in close proximity to the pipeline network. Specifically, I find that a 10% reduction in pipeline distance produces an additional 6.4% increase in likelihood of conversion for a generating unit subject to MATS. I examine these incentives more thoroughly within a structural framework, which allows me to isolate the contribution of pipeline proximity to a coal plant's fixed cost of conversion. This dynamic model of the plant's decision to convert allows me to estimate the impact of infrastructure as a potential policy lever for clean energy use, even in absence of MATS. Model estimates demonstrate large-scale benefits from pipeline subsidies that amount to nearly one-third the benefits of federal mandates through MATS. These results suggest that pipeline subsidies alone can lead to present value external benefits of up to $10.1 billion in emissions reductions from coal-to-natural gas conversions. Additional calculations estimate a marginal value of the pipeline infrastructure at $2.6 million per-mile.

2. ARMS RACE DYNAMICS IN VEHICLE PURCHASES

2.1 Introduction

The average size of the U.S. vehicle fleet has risen steadily over the years. According to the National Highway Traffic Safety Administration (NHTSA), passenger cars have observed a 12.5% increase in weight, and light trucks a 26% in weight from model year 1984 to 2004. This increase in vehicle sizes can be seen in Figure A.1. With vehicles on the road contributing to nearly one-fifth of carbon emissions, and the negative relationship between vehicle size and fuel efficiency, this growth in fleet size has added significantly to the increase in carbon emissions over the years, and has impeded much of the progress in fleet fuel efficiency.

The increased size of the vehicle fleet is driven by multiple factors. Consumers throughout most of this period have experienced growths in income, so an increased willingness to pay for vehicle size might explain a portion of the growth in fleet size. The increase in fleet size may also be driven by relatively low real prices of fuel, reducing the cost of driving these large cars. This paper examines another factor contributing to the growth in fleet size: consumers strategically responding to heavy cars on the road, thereby increasing their purchases of sport utility vehicles (SUVs) and pickup trucks, relative to passenger cars, in order to “protect" themselves from other vehicles on the road. Anderson and Auffhammer (2014) demonstrate that large weight disparities in two-car collisions significantly increase the likelihood of a fatality occurring. Their work documents the fact that “pounds kill," however, no research has shown whether consumers incorporate the fact that “pounds kill" into their purchase decisions. This paper seeks to answer the question of whether consumers know that “pounds kill" and purchase accordingly.

Households may purchase larger vehicles due to preferences for safety. The same household may prefer a larger vehicle if the cars owned by other households in the area are larger. The literature on the vehicle "arms race" (e.g., White, 2004; Li, 2012) discusses these preferences for safety, wherein consumers buy increasingly large vehicles to protect themselves from other cars on the road. This type of contagion effect, inherent in a vehicle arms race, produces higher emissions and increased risk on the road as vehicles increase in size. Furthermore, these responses create feedback loops that cause the fleet to be heavier than it would otherwise be, absent the strategic motive by households to protect themselves.

The traditional view is that heavier vehicles provide more safety when experiencing an accident (Crandall and Graham, 1989). However, consumers who protect themselves by purchasing pickup trucks and SUVs create an increased risk on the road to other drivers. It is well-documented that large vehicles create an increased risk of fatality to the opposing driver when involved in an accident. For example, Jacobsen (2013) finds that a 1,000 pound increase in the weight of a vehicle involved in an accident increases the number of fatalities in other vehicles by about 46 percent, but reduces own risk by 53 percent. Anderson and Auffhammer (2014) find, conditional on the occurrence of an accident, a 1,000 pound increase in vehicle weight is associated with a 47% increase in the baseline fatality probability. Evans (2001) finds that the effects of adding mass, in the form of a passenger, adds to the increased risk of fatality in head on collisions. He finds that adding a passenger to a car leads to a 7.5% reduction in driver's risk of fatality, while increasing the risk of fatality to the other driver by 8.1%. Bento et al. (2017), similar to Anderson and Auffhammer (2014), examine the relationship between weight dispersion and fatalities in accidents, but in the context of CAFE. They find that, though CAFE increased the national fleet's weight dispersion, the resulting decrease in average weight produced an overall reduction in fatalities.

These types of safety attributes, and the extent to which consumers understand the relationship between size and safety, may create strong preferences for *relatively* large vehicles. To my knowledge, no prior research has provided causal evidence that households respond to heavier vehicles they encounter on the road by purchasing larger cars. I examine this relationship in this paper.

Furthermore, I show that these arms race dynamics are a significant determinant of the increasing fleet size. Finally, I demonstrate how Pigouvian taxes, when accounting for an arms race, may counteract these types of preferences.

The types of incentives uncovered in this paper have significant externalities. For example, because larger cars tend to be less fuel efficient, increased demand for heavier vehicles has a conflicting effect on policies seeking to improve vehicle fuel economy. CAFE standards aim to reduce energy consumption by increasing the fuel economy of the U.S. vehicle fleet through manufacturer targeted minimum fuel efficiency requirements. The CAFE standards have had positive effects on improving average vehicle miles-per-gallon (MPG), however, the vehicle arms race creates an opposing dynamic, reducing the overall potential of CAFE.

Accounting for the dynamic preferences for weight, inherent in an arms race, is especially important when evaluating Pigouvian taxes. For example, a tax on gasoline which takes into account consumers' static preferences for large vehicles ignores how these preferences might change in response to other cars on the road. In an arms race, consumers favor large cars increasingly more as the fleet size around them grows. Not accounting for these dynamic preferences can lead to mis-estimation of a tax. Thus, credible identification of arms race preferences is an important step in improving these policies.

Industry analysts have conjectured that an arms race effect has played a major role in the growth in size of the vehicle fleet presented in Figure A.1. The logic is captured in a New York Times article on the "SUV Arms Race," quoting a consumer saying, "I'm sure the world would be a better place if we all drove Minis, but with so many crazy drivers on the road, I want the biggest vehicle I can get."

Though analysts have speculated about the apparent "arms race" in the U.S., there has been no credible evidence that this type of behavior actually exists. A primary reason is the difficulty in identifying such an effect using naturally occurring field data. The thought experiment would be to randomly assign households in some region with large vehicles and another region with small vehicles, and observe the weights of future car purchases. Such an experiment would isolate confounding factors such as common shocks and unobserved characteristics that are common across households, and also reveal preferences for vehicle size. Random assignment would also mitigate concerns of simultaneity bias; we would be able to properly identify the causal direction of the treated households impact on other households. However, such large scale experiments are often infeasible in practice.

In this paper, I exploit a source of quasi-random variaton in treating households with a salient shock to consumers' preceptions about vehicle safety. Specifically, I compare the behavior of *next-door* neighbors of households that were involved in a serious accident. To explain, consider a two-car collision, in which the driver of one car dies and the driver of the other car survives. This accident is a salient information shock to *acquaintances* of the drivers involved in the accident, and may impact the future vehicle purchase behavior of those acquaintances. To proxy for unobserved acquaintances in these crash victims' networks, I compare the subsequent vehicle purchase behavior of the next-door neighbors of the drivers who died to the purchase behavior of the next-door neighbors of the drivers who survived.

I use a difference-in-differences empirical strategy to measure the impact of this salient preference shock. To implement this strategy, I use a rich administrative database of vehicles registered in Texas from 2004-2010, matched to a census of all vehicle accidents in Texas. Using a panel of household-level vehicle purchase data, I study the changes in the weights of the neighbors' vehicle purchases after the accident. I find that prior to the accident, the next-door neighbors purchased vehicles with extremely similar weights. However, after the accident, the next-door neighbors of the driver who died in the accident owned vehicles that were 20-30 pounds heavier than the next-door neighbors of the driver who survived (see Figure A.2). This amounts to roughly one in ten treated households upgrading from a sedan to a mid-sized SUV.

I exploit another source of quasi-random variation to verify these vehicle purchase dynamics. I study two-car accidents, where it has been shown that increases in weight of the striking vehicle will increase the fatality risk of passengers in the smaller vehicle (Anderson and Auffhammer, 2014). I show that when a victim is quasi-randomly hit by a heavy vehicle, and the victim is driving a smaller vehicle, the next-door neighbor responds by purchasing a larger car, following the accident. This exercise provides further evidence that a salient shock to preferences for safety results in higher demand for large vehicles.

I consider these responses by consumers to be evidence of information shocks to their perceptions of vehicle safety. This is primarily due to the salient information that these households are treated with. That is, a household now becomes aware of an acquaintance's fatality in an accident. This type of information evidently creates an adjustment in preferences for vehicle attributes related to safety. The primary determinant of these changes in preferences is plausibly due to a change in beliefs about the safety of their current vehicles.

In this paper, I do not take a stand on potential behavioral mechanisms that could be driving these changes in preferences. One can speculate that these responses may be a result of projection bias (Lowenstein et al., 2003; Conlin et al., 2007; Busse et al., 2014), anchoring heuristics (Slovic, 1967; Tversky and Kahneman, 1974; Crandall and Graham, 1989; Furnham, 2011), or other assimilative biases. Identifying any behavioral mechanisms that may be driving these responses is not feasible given the information that I have. Furthermore, I cannot reject the possibility that consumers are simply rationally updating beliefs from incorrect priors.

This effect of quasi-random shifts in preferences for safety provides credible evidence that arms race dynamics affect vehicle purchases. It is important to understand how these preferences will impact the evolution of the vehicle fleet over time, and how policies might operate to internalize an arms race. To do so, I use a discrete choice model to directly estimate revealed preferences for vehicle size, as a function of the vehicle fleet around a consumer. This model allows me to simulate the long-term effects of an arms race feedback loop under various counterfactual policy schemes.

I estimate a discrete choice logit model of vehicle purchase that allows tastes for vehicle weight to vary based on the distribution of vehicle sizes in a consumer's local area. This model directly estimates consumer preferences, and thus, I will be able to place a dollar value on *relative* vehicle weight. That is, I consider how shifts in vehicle sizes within an area change how consumers value their own vehicle size. In order to estimate this model, I need to account for the possibility that unobservered preferences for weight are correlated with the weight distribution in a consumer's area. To do so, I devise a unique instrument based on the variation in past gasoline prices, internalized by consumers at past times of purchase, as well as beliefs about future gasoline prices (Anderson et al., 2012). Results indicate that a consumer's value of an additional pound of weight increases by $0.02 following a one pound increase in average county vehicle size. This implies a consumer would value a Jeep Cherokee (SUV) about $3,200 more than a Honda Accord (sedan), when treated with a county driving all Ford Explorers (heavy county), compared to a county driving all Ford Fusions (light county).

Finally, I use the model to calculate the impact of counterfactual events, such as a sustained period of very low, or very high gas prices. This exercise provides insight into the impact of static Pigouvian taxes in an arms race. I find that, with sufficiently high gas prices, an arms race can be reversed. That is, if consumers find it too costly to increase vehicle size, weight distributions shift downward, and thus, consumers are less endangered by the relatively lighter vehicles on the road. With high gas prices sustained for long periods of time, consumers have a lower likelihood of being harmed by the previously purchased, smaller vehicles on the road, and therefore benefit from reducing vehicle sizes further. This creates a downward spiral in vehicle sizes with large environmental benefits. However, when gas prices are low, the arms race feedback loop is exacerbated, which leads to substantial increases in fleet size. Since consumers internalize their preferences for weight and preferences for safety—which depend on the size of their local fleet—the optimal Pigouvian tax to internalize an arms race will ultimately depend on the local weight distributions.

These exercises illustrate the importance that optimal tax calculations take into account consumers' dynamic preferences for vehicle size. Taxing schemes may differ based on the policy objective; however, I examine a tax to fully internalize an arms race feedback loop. Simulation results indicate, to fully internalize an arms race, in many instances, it is most efficient to stimulate a reduction in vehicle size with an initial tax rate increase, kick-starting a reverse arms race, then following up with a tax reduction. The initial tax increase serves to replace many of the large cars on the road with smaller cars, due to the increased price of driving large vehicles. This ultimately shuts off much of the feedback loop, as consumers grow less concerned for safety when the vehicles around them drop in size. Once the arms race has been neutralized, estimates indicate it is efficient for taxes to drop.

The evidence of a vehicle arms race I present in this paper provides new evidence of a major factor driving the increases in fleet size over the years. These preferences exist even in the presence of government standards, such as CAFE, that aim to improve fuel efficiency, and thus, place downward pressure on the size of the fleet. These types of strategic responses by consumers therefore identify a new dynamic which policy efforts must account for when designing policy.

No policy would be able to eliminate the type of preferences that I uncover in this paper. However, internalizing some of the externalities produced in an arms race can significantly weaken the growth in fleet size produced by this feedback loop. This paper specifically looks at a gas tax as a potential policy lever to moderate the growth in vehicle size. When it becomes too costly to drive trucks and SUVs, consumers will respond by purchasing smaller, more fuel efficient cars, which weakens the best response mechanisms of subsequent purchasers. Though not directly examined in this paper, a feebate—some combination of fees and rebates—implemented in the market for new vehicles may also produce a more efficient outcome. It is important to note, that any policy addressing the costly externalities produced by arms race preferences, must take into account the feedback loops generated when consumers respond in purchases to other vehicles on the road.

The paper proceeds as follows: In Section 2.2, I lay out the groundwork for my empirical identification strategy. In Section 2.3, I discuss the data and dataset construction. In Section 2.4, I present my primary estimation results of the effect of neighbor's fatality and a series of robustness checks. In Section 2.5, I provide further quasi-experimental evidence of an arms race by examining the additional impact of large vehicles in vehicle accidents. In Section 2.6, I introduce and estimate a model of consumer choice in an arms race. In Section 2.7, I use the estimated model to calculate the magnitude of an arms race in different Pigouvian scenarios. In Section 2.8, I conclude my paper with a brief discussion of the implications of my findings.

2.2 Empirical Framework

This paper has two specific research questions that will address two distinct aspects of a vehicle arms race. First, I examine how consumer preferences for vehicle safety can drive demand for large cars. That is, I study whether consumers behave in a manner consistent with an arms race, or have arms race-*style* preferences. To identify these responses, I exploit vehicle accidents as an exogenous shifter of preferences for safety. Second, I examine how consumer preferences for vehicle size change as a function of the size of vehicles on the road, and policies that can internalize arms race externalities. This approach will require a model of vehicle choice, where preferences depend on local distributions of vehicle sizes.

To study consumer preferences for safety in an arms race, I narrow my focus to salient outcomes which large vehicles produce in accidents. Specifically, I study *next-door* neighbors of individuals involved in fatal accidents. Among those involved in an accident with a fatality, some will live and others will die. Thus, some of the next-door neighbors will be "treated" with the knowledge that someone who they likely knew died in a vehicle accident. I view this treatment as an information shock to consumers' perception of vehicle safety, which is most likely influenced by the size of the vehicles involved in the accident.

To illustrate how this shock to perceived preferences for safety might impact consumer behavior, consider an additive utility function, where consumers place value on the safety associated with a certain vehicle's weight. In the spirit of Anderson and Auffhammer (2014), suppose this value enters the utility function negatively in the form of a consumer's belief about the probability of dying in an accident, conditional on the weight of that vehicle and the distribution of vehicle sizes in the population, $f(\cdot)$. Let $h_i(w_j, w)$ be a function *proportional* to consumer i's beliefs about the probability of dying in a vehicle, j, of weight w_j, when colliding with another vehicle of weight w (i.e., $Pr(\textit{die in an accident}|w_j, w)$ enters utility linearly). Suppose consumer i chooses a vehicle subject to a budget constraint, with the price of a numeraire good normalized to one. Let i's indirect utility function for vehicle j be described by the following.

$$u_{ij} = \alpha_i(y_i - p_j) + x_j\beta_i - \int h_i(w_j, w)f(w)dw + \varepsilon_{ij} \tag{2.1}$$

where α_i describes consumer i's marginal utility of income, y_i is i's income, p_j is the price for car j, β_i are preference parameters on various attributes, x_j (may include vehicle size), and ε_{ij} are i's unobserved preferences for vehicle j. Suppose consumer i updates their beliefs of $h_i(\cdot)$ following the death of a neighbor in an accident to $h_i(w_j, w \,|\, neighbor\_died_i)$. That is, a neighbor's fatality increases the saliency of the role that weights play with respect to safety.

For example, in the extreme case, suppose prior to the accident beliefs enter utility as a constant, or flat in weight (i.e., $\partial h_i(w_j, \cdot)/\partial w_j = 0$). Following the neighbor's death, consumer i updates beliefs such that the probability of dying is lower if one is driving a heavier vehicle (i.e., $\partial h_i(w_j, \cdot \,|\, neighbor\_died_i)/\partial w_j < 0$). Not directly examined in this paper, this updating may be a result of projection bias, incorrect priors, or some other behavioral mechanism. The extent to which the consumer updates these beliefs, conditional on a neighbor dying, governs the extent to which the consumer changes their preferences for safety.

2.2.1 Quasi-experimental Evidence of Purchase Response to Shocks to Preferences for Safety

Evidence of an increased preference for safety following a neighbor's death can be directly observed through the consumers' revealed preferences for weight. The extent to which consumers adjust these preferences according to the distribution of vehicle sizes on the road, $w \sim f(\cdot)$, (as in Equation 2.1) is examined directly in Section 2.6. The key focus of the first part of this paper, however, is in isolating these shocks to consumers' preference for safety.

My empirical strategy to isolate these shocks to preferences for safety is to implement a difference-in-differences estimator of the treatment effect of neighbor's fatality on household choice of vehicle size. The key identifying assumption is that subsequent vehicle purchasing behavior of the treated group would have followed a parallel pattern of the control group, absent treatment. I test if the two groups are parallel, absent treatment, by testing if trends in outcomes were similar prior to the accident. Specifically, I estimate the following fixed effects regression:

$$weight_{it} = \alpha_i + \lambda_t + \beta \cdot neighbor\_died_{it} + \varepsilon_{it} \tag{2.2}$$

where $weight_{it}$ is the average curb weight of household i's fleet in month-of-sample t, and $neighbor\_died_{it}$ is an indicator which equals one if household i died in a vehicle accident in some period $s \leq t$, and zero otherwise. The parameter of interest is β which estimates the average treatment effect on the treated (ATET). As I observe multiple purchases from the majority of households in my sample, I control for household level and month-of-sample fixed effects, α_i and λ_t, respectively.

I will estimate Equation 3.2 using, as my control group, households neighboring an individual who was involved in a fatal accident (at least one crash victim died), but who survived it. The intuition is that this group of individuals will likely behave similar to the treatment group on average. It is likely that the control group had larger vehicles on average than those of the treatment group. This might be the case because of the correlation between vehicle size and likelihood of fatality in an accident, as well as the correlation in characteristics among neighbors. This alone does not violate the identifying assumption. It is only required that these differences remain fixed over time, prior to the accident.

Using the neighbors of the *survivors* as the control group could lead to smaller estimates than if the control group included households who had no neighbor experience a car accident at all. This could be the case if the control group was affected (in the same direction) by their neighbor's accident too, even though that neighbor survived. However, using the *survivors* as the control group should provide robust estimates because of the presumed similarity they have with the treatment group, but with a distinguishing feature (i.e., surviving) that will help disentangle the effects of a neighbor's death.[1]

2.3 Data

The data for this paper are derived from vehicle registrations and accident records. For vehicle accidents, I use information on accidents involving at least one fatality. These data come from the

[1]One could argue that using the *survivors* as my control group may make the effect seem larger than it actually is. This could be the case if the neighbors of these survivors decided to purchase smaller vehicles following the accident. Therefore, to lead to an over-sized estimate of the effect on the treatment group, the control households would need to be unrealistically conscious of the accidents and willing to reduce the size of their vehicle to "save" others' lives. I view this social utility hypothesis as unlikely to be the case, and thus, posit that my estimates will be a lower-bound of the true effect of treatment. With that said, in the Online Appendix B.4 I explore an alternative control group and demonstrate that my choice of control group does not significantly alter my estimates.

Fatality Analysis Reporting System (FARS) and are supplemented with a Texas A&M Transportation Institute (TTI) dataset, which is a census of police accident reports for *all* Texas accidents. All accidents data are observed over the years 2006 through 2010. For household vehicle ownership, I use confidential administrative records maintained by the Texas Department of Motor Vehicles (DMV). These data include all registration records for 2004 through 2010.

In this paper, a majority of the analysis will be conducted on neighbors of crash victims. This requires identifying the victims of these crashes in the registrations data. The DMV data provide the exact addresses of these households, given their last vehicle registration. Once I have these addresses, I can locate physical *next-door* neighbors, by finding households located within the nearest distance, and on the same street.

For each household neighboring a crash victim, I decode the vehicle identification numbers (VIN) for their entire fleet using a database obtained from DataOne Software. These data allow me to merge in various characteristics of each vehicle, such as make, model, year, weight and MSRP. The majority of the analysis in this paper focuses on the size of the households' vehicle fleets.

Due to an absence of VIN or household-level information in the FARS dataset, I make use of additional information in the TTI data to "fuzzy" match fatal vehicle accident victims to households in the DMV dataset. This process involves two layers of "fuzzy" matching, in which I first match key TTI information to FARS observations, and then match fatal accident victims to vehicle owners in the DMV data. Combining these two datasets provides me with (non-unique) 12-digit substrings of 17-digit VINs, dates of crashes (which provide information on ownership times), as well as the drivers' first and last name. I proceed by using this information to "fuzzy" match to ownership records in my DMV data.

My final dataset includes about 46% of the original FARS dataset. For these households, I now have information from the DMV dataset such as vehicle ownership and addresses. I locate the neighbors of these households, and their corresponding ownership histories. It is important to recognize that this matching procedure will most likely lead to some measurement error in the treatment variable. That is, I may incorrectly identify a household as being involved in a fatal accident. Estimated treatment effects may, therefore, be attenuated.

For the purpose of estimating the effect of a neighbor's fatality on nearby neighbors, I will restrict the set of neighbors to those located on the same street as the individual involved in the accident, and within 100 meters. I primarily focus on the two nearest neighbors, but will show robustness through other samples.

2.3.1 Summary Statistics

Summary of statistics are presented in Table A.11. Each column in the table corresponds to a different sample containing a different number of nearest neighbors. These are the k closest households neighboring the individual who experienced a vehicle accident. Note that the sample size is not exactly proportional to the number of nearest neighbors included. This is either the result of missing data, small streets, or neighborhoods with spaced out houses (as only houses on the same street and within 100 meters are included in my dataset).

About 56% of households in my sample are treated with a neighbor fatality. The average vehicle curb weight in this sample is about 3,700 pounds. This is about the mass of a medium-sized SUV. About 67% of households in the sample own at least one passenger car, as opposed to owning *all* light trucks.

Another key characteristic of this sample is that the average household purchases just over two vehicles over the course of the seven years of my registrations data (2004-2010). Because of these quite infrequent purchases, any adjustment to household vehicle characteristics will most likely be a slow and gradual process. Note that the fixed-effects estimators in this paper will fully absorb one-car purchase households, and will only exploit the variation coming from households who purchase multiple vehicles within the time-frame of my data.

2.4 Quasi-Experimental Evidence of Arms Race Dynamics

2.4.1 Effect of a Neighbor's Fatality

In this section, I look at the effects of these information shocks on various outcomes. Specifically, I test whether a neighbor's fatality in a vehicle accident has an effect on the revealed preferences for vehicle size.

I examine consumer behavior at the household-level, on a monthly-level panel and estimate Equation 3.2, where the outcome is a snapshot of the size of the household fleet at a given month in time. Note that the characteristics of a household fleet will only change when the household purchases a new vehicle. For these new purchases I observe the exact date, however for simplicity, I collapse these into monthly observations.

Evidence of a treatment effect on neighbor's death can be seen in Figure A.2. This figure plots the average vehicle size for the treatment and control groups on months away from treatment. I have included only the two nearest next-door neighbors in this figure, though I check the robustness across different samples of neighbors in Section 2.4.2. As expected, the neighbors of those who died tend to track well against the neighbors of the survivors in the periods preceding treatment. At the treatment period, their paths begin to diverge. This figure illustrates that the treatment group begins to increase the relative size of their vehicles following neighbor's death.

It is apparent that the control group breaks trend around the time of the accident. Note that this should not bias my estimates, so long as my identification strategy is valid—that is, as long as the treatment group would have continued to follow the same trend with the control group, absent a neighbor's fatality. In Online Appendix B.2, I show that this trend break is entirely generated by time specific factors[2], and thus should not bias my estimates. I also examine an alternative control group in Online Appendix B.4 and confirm similar results.

The parallel paths of the treatment and control groups in Figure A.2 provide strong evidence that I can accurately predict the treatment group's counterfactual path, absent treatment. Table A.2 presents the least squares estimated treatment effect on household average vehicle weight. The sample includes the nearest two neighbors to the individual involved in the accident.

Column 1 estimates the treatment effect using the traditional difference-in-differences estimator, which controls for a post-accident indicator, as well as a treatment group indicator. In Column 2, I control for time-invariant census tract-level demographic variables, such as median income, percent white, percent black, and median age. Column 3 includes month-of-sample fixed effects, and in Column 4, I estimate Equation 3.2. This includes a full set of household and time fixed effects. Note that these estimates fully absorb any variation coming from households who have no new purchases during the time-frame of my sample. This may explain the decrease in the size of these point estimates after including household-level fixed effects. In the last column, I include a county-specific linear time trend, allowing each county to follow differing trends.

These estimates indicate that households residing near a neighbor who died in an accident, increase their average vehicle size by about 20 to 30 pounds following a neighbor's fatality. This implies that, for a two-vehicle household purchasing only one new car in the post period, this vehicle should on average be about 40 pounds heavier. This is equivalent to a 400 pound upgrade for one in ten two-car neighbors—roughly the same as a switch from a sedan to a mid-sized SUV.

These estimates suggest that a neighbor's fatality in a vehicle accident makes vehicle weight salient to these households. This type of behavior is consistent with preferences in a vehicle arms race,

[2]The median crash time in my data is June 2008. This was a period experiencing very high gas prices—one factor that may have contributed to this trend break.

where the neighbor's death in a car crash serves as a quasi-random shock to preferences for safety.

2.4.2 Different Neighbor Samples

Next I estimate these treatment effects using different samples of nearest neighbors. In each specification, I include the nearest one through k nearest next-door neighbors (on the same street), as listed in Table A.3. These estimates serve as a falsification, as one would expect the estimates to become attenuated as I add neighbors who live further away—who are less likely to be acquaintances with the crash victims.

Examining the results in Table A.3, the two neighbor sample appears to capture the largest effect in terms of household vehicle weight (though not statically different from the three neighbor sample). These neighbors are more likely to be acquaintances with the crash victims than neighbors further away. The estimates show a gradual decay in size following the second neighbor. The effect of neighbor's fatality is no longer significant by the ten nearest neighbors sample.

2.4.3 Dynamic Treatment Effects

Next, I estimate a fully dynamic version of Equation 3.2. In the regression, I include six leads and twenty-four lagged indicators of treatment. The final lagged term is an indicator for 24 or more months post-fatality—it reflects the remaining time-span of the data. The leading terms serve as placebos, as I should not expect future fatalities to impact current household vehicle sizes. This formally tests the parallel trends assumption, as displayed in Figure A.2. The dynamic treatment effects presented in Figure A.3 also illustrate the gradual effects of neighbor's fatality on the average weight of the households' vehicle stock overtime. These estimates indicate that, a year out from the treatment period, treated households are driving vehicles roughly 20–25 pounds heavier.

2.4.4 Placebo Tests

I further check the robustness of the findings in Table A.2 by estimating the treatment effect for a series of placebo estimates. This procedure is similar to the one outlined by Bertrand et al. (2004).

Using the same sample as in Table A.2 (two neighbor sample), I randomly assign each of 15,723 households to the treatment or control group, with the same frequency of treatment as the empirical probabilities. I randomly assign a treatment period to each household from the periods in which the household was observed. For most households, I observe their vehicle stock for the duration of the accidents data, which begins in 2006 and ends in 2010. I proceed by estimating Equation 3.2 for these placebo treatments. I repeat this process 1,000 times and record the estimates for each.

This process leaves me with a distribution of placebo estimates. In a manner similar to that of synthetic control methods (e.g., Abadie et al., 2010), I calculate one-tailed p-values for my estimates in Table A.2 from the resulting distribution. A kernel distribution of the placebo estimates is presented in Figure A.4.

As expected, this distribution is centered on zero, which indicates a mean zero effect for a randomly assigned placebo treatment. My estimate for the effect of neighbor's fatality on average household vehicle weight ranks at the top of the distribution, with a corresponding p-value of 0.3%. This implies that less than half a percent of placebo treatements resulted in a greater estimate than my estimate of 19.95 (from Table A.2), suggesting it is highly unlikely that an effect of this size was estimated by chance.

2.5 Spillover Effects of Crash Vehicle Weights

In this section, I provide further evidence of arms race-*style* preferences by examining how households neighboring crash victims respond indirectly to the size of vehicles involved in the accidents. The idea is that a larger striking (or opposing) vehicle, and a smaller vehicle driven by a neighbor, produces more severe and salient outcomes (Anderson and Auffhammer, 2014). The household neighboring the victim then directly responds to these salient outcomes—a direct function of the crash weights.

Identification of indirect responses to the size of the vehicles involved in these accidents come from the as-good-as-random timing of the accidents. Neighboring households may directly respond to each other's vehicle size by means of social norms. I will not be able to identify these *baseline* effects, but rather, identify the *additional* response to a victim's vehicle size following the accident. I estimate these indirect responses from a household to both the neighbor's and the opposing vehicle's weight in a two-car collision.

To identify the spillover effects of the vehicle sizes involved in an accident, I estimate the following equation:

$$\begin{aligned} weight_{it} = \alpha_i + \lambda_t + \gamma \cdot post\_crash_{it} + \beta_1 \cdot weight_i^{nbr} \times post\_crash_{it} \\ + \beta_2 \cdot weight_i^{opp} \times post\_crash_{it} + \varepsilon_{it} \end{aligned} \tag{2.3}$$

where α_i and λ_t are household and month-of-sample fixed effects, respectively. Household fixed effects will absorb any baseline sizes of the two vehicles involved in the accident that the household may respond to prior to the accident. In some specifications I will simply control for these baseline weights, rather than household fixed effects (which absorb all one-purchase households in my data). $post\_crash_{it}$ is an indicator for i's neighbor incurring a vehicle accident at some time $s \geq t$, and ε_{it} is the unobserved error. β_1 estimates the additional response to the weight of the neighbor's vehicle involved in the accident, $weight_i^{nbr}$. β_2 estimates the additional response to the weight of the opposing vehicle's weight, $weight_i^{opp}$, which hit the neighbor.

Note that $\beta_1 < 0$ implies that a smaller neighbor's vehicle involved in the accident on average produces a more severe outcome, leading the household to purchase a heavier vehicle. An estimate of $\beta_2 > 0$ implies that a larger vehicle hitting the neighbor's car creates a more severe outcome, leading the household to increase its vehicle size.

2.5.1 Results and Robustness

The least squares estimates of Equation 2.3 are presented in Table A.4. These estimates report the effect of each 1,000 pounds of the crash vehicles' weights. The first two columns estimate the effect of victims' vehicle weights on neighbor fatality, using a linear probability model. This is simply the cross-section of vehicle-by-accidents. The results indicate that a 1,000 pound increase (decrease) in opposing (neighbor's) vehicle weight increases the probability of neighbor fatality by about 6-7%. In the second column, I include census tract-level demographics, such as median income, median age, percent white, and percent black.

Because the weight effects not only risk of fatality, but also the severity of accident (and potentially other salient factors), using these crash weights as an instrument for neighbor fatality would violate the exclusion restriction. Therefore, inference should only be made from the reduced form effects of crash weights on neighbor of victim weight. These estimates are presented in Columns 3-7 of Table A.4.

There is a similarity in the estimates reported in Columns 1-2 of Table A.4 and the estimates reported by Anderson and Auffhammer (2014). They find that an additional 1,000 pounds of the striking (or opposing) vehicle generates about a 0.109 percentage point increase in risk of fatality. Note that I should not expect to get the same estimate, as my sample conditions on accidents involving at least one fatality, whereas Anderson and Auffhammer's dataset also include accidents involving no fatalities.

The coefficients in Table A.4 were estimated using the two-neighbor sample. The results for the effect of opposing vehicle size indicate a statistically significant estimate of about 16-18 pounds. That is, a 1,000 pound increase in the opposing vehicle's weight causes household i to increase their average fleet weight by 16-18 pounds.

The effect of the neighbor's vehicle size represents the converse of the opposing vehicle's size. That is, the *smaller* the neighbor's car, the larger the response. Columns 3-4 illustrate a similar-in-magnitude response as the impact of oppsing vehicle. The inclusion of household fixed effects aborb much of this effect. Though estimates with and without household fixed effects are not statistically different from each other, conditioning on household fixed effects seems to eliminate statistical significance.

The reduced-form estimates of opposing vehicle weight offer corroborating evidence of arms race dynamics using another plausible source of quasi-random variation. The opposing vehicle is matched to the neighbor as-good-as-randomly, and the likelihood that these two victims were acquaintances is therefore small. I am also exploiting the as-good-as-random timing of the accident for these estimates—the only source of (plausibly) random variation in estimating the effect of the neighbor's fleet weight, post-crash.

I present the estimates for samples containing differing numbers of nearest next-door neighbors. These results are presented in Table A.5. Similar to previous results, these estimates become attenuated as I include neighbors further away. However, the estimates for opposing vehicle do not fully converge to zero in the 10 neighbor sample. This result differs from the estimates in Table A.3, where the estimates on weight converge to zero quickly.

In the event that these estimates remain positive for neighbors very far away, one should be concerned with the potential influence of confounders, such as common shocks. To mitigate these concerns, I could look at neighbors located on a street behind the household, as these households are less likely to be acquaintances with the crash victims. The process of finding neighbors a block over is very tedious. The simpler approach is to randomly assign neighbors to opposing vehicle weights from the sample. This is similar to the permutation approach in Section 2.4.4, except performed in one dimension. That is, I maintain the same crash time (the true crash time), which is a source of random variation, but randomly re-assign a crash vehicle pair from the sample of two-car crashes. If I am picking up confounding factors common across all vehicles, I would expect to see effects not only from the opposing vehicle hitting the neighbor, but also other opposing vehicles in the sample.

I randomly draw, with replacement, victims of vehicle accidents and re-assign them to households. That is, I re-assign neighbors, collecting the weights from both vehicles in the accident. I make this re-assignment 1,000 times, estimating the coefficients in each permutation, and plot the distribution of these placebo estimates. Using this approach resulted in a one-tailed p-value of 0.082 for the effect of neighbor's weight, and 0.001 for the opposing vehicle. The distribution of the placebo estimates are presented in Figure A.5. These results suggest that it is highly unlikely that the estimates in Tables A.4 and A.5 are the outcome of common shocks.

2.6 Consumer Preferences in a Vehicle Arms Race

In this section, I develop a model of consumer choice in a vehicle arms race. By directly estimating consumer preferences, I evaluate in dollar terms how much consumers value relative weight. An additional benefit of estimating this model is that it will allow me to estimate counterfactual scenarios of policies aimed at "undoing" the arms race feedback loop. Specifically, I examine a targeted Pigouvian tax on gasoline.

2.6.1 The Model

I develop a simple static discrete choice model of vehicle purchases in order to evaluate the long-run effects of a vehicle arms race. One may be concerned with the precision of the parameter estimates when assuming static demand in the context of durable goods. It is true in the vehicle market that consumers may substitute consumption intertemporally. Their purchase decisions may depend not only on current prices, but also expected future prices. Ignoring dynamics could, therefore, bias price elasticity estimates significantly (e.g., Hendel and Nevo, 2004). Assuming static demand may be a problem, particularly in an arms race, where vehicle size decisions may depend on future expectations of the size of vehicles on the road. In this context, one would presume that static estimates would underestimate a consumer's choice of vehicle size. Parameter estimates under static demand assumptions, therefore, must be interpreted with caution. Nevertheless, this simple model allows me to illustrate the long-run effects of arms race purchase behavior.

I specify indirect utility for some vehicle j to be a linear function of the price and product attributes of that vehicle. I model preferences for vehicle size in an arms race by allowing utility for weight to vary based on county-level weight distributions. Let the following be consumer i's indirect utility in county c for vehicle j:

$$\begin{aligned} u_{icj} &= \alpha_i(y_i - p_j) + x_j\beta_i + \bar{h}_{ic}(w_j) + \varepsilon_{icj} \\ &= \nu_{icj} + \varepsilon_{icj} \end{aligned} \tag{2.4}$$

where I omit time, t, subscript for simplicity. α_i is the consumer-specific marginal utility of income, y_i is consumer i's income, and p_j is the price of car j. Various product attributes are included in vector, x_j, with preference parameters, β_i. ε_{icj} captures the idiosyncratic tastes for vehicle j.[3]

I specify preferences for weight to enter utility according to a consumer-county specific function $\bar{h}_{ic}(\cdot)$. This function allows preferences for vehicle size to vary based on the distribution of weights in the county. To explain, suppose consumer i's preferences for vehicle weight, w_j, vary depending on the size of another vehicle on the road, w. Let these preferences be dictated by the function $h_i(w_j, w)$. For example, Anderson and Auffhammer (2014) propose that consumers' preferences for weight vary depending on the probability of dying in a two-vehicle accident, conditional on the weight disparity between the two vehicles involved in the accident. That is, $h_i(w_j, w)$ may be a

[3]In this simple model, a household chooses to purchase a vehicle as a function of that vehicle's price, characteristics, and the weight of the nearby fleet. In recent work, Archsmith et al. (2017) study another important and nuanced form of decision-making in which a household's choice of a new car depends on attributes of other vehicles in that household's portfolio.

function proportional to $Pr(\textit{die in an accident}|w_j - w)$.

Because consumer preferences for vehicle size vary not just on one other vehicle on the road, but plausibly all the vehicles on the road, preferences enter as the expectation of $h_i(w_j, w)$ conditional on the distribution of vehicle sizes in an area. For simplicity, I specify preferences for vehicle sizes as a function of the county-level weight distribution. That is, consumer i's preferences for weight, w_j, in county c, are dictated by the following function:

$$\bar{h}_c(w_j) = \int h(w_j, w) f_c(w) dw$$

where $f_c(\cdot)$ is county c's weight probability density function.

I specify weight preferences at the *county*-level because most consumers drive within close proximity to home. This is also consistent with Anderson and Auffhammer's specification, which depends on the probability of fatality in a car accident, since most car accidents occur close to home. In fact, a 2001 Progressive Insurance Survey found that 52 percent of car crashes occur within five miles or less from home, and 77 percent occur within 15 miles from home.

I assume that consumers purchase the vehicle that maximizes their expected utility. When unobserved preferences, ε_{icj}, are distributed independent and indentically from the Weibull (or type I extreme value) distribution, the choice probabilities have a closed form solution. That is, given a set of J alternatives, the probability that the consumer chooses car j is:

$$s_{icj} = \frac{exp(v_{icj})}{\sum_{k \in J} exp(v_{ick})} \tag{2.5}$$

2.6.2 Estimation

To estimate the model, I use the same registrations data from earlier parts of this paper. For computational simplification, I restrict the sample to purchases of new model year vehicles made by households residing in six counties in the Houston Metropolitan Statistical Area (MSA).[4] These counties provide significant variation in vehicle sizes, as preferences presumably vary over these urban and rural areas.

In estimating preferences for weight that vary based on the county weight distribution, I do not impose a non-linear structure on $h(\cdot)$. That is, I do not restrict $h(\cdot,\cdot)$ to the probability of dying in an accident, as consumers may respond to surrounding weights by other aversion metrics. For simplicity, I specify $h(\cdot,\cdot)$ to be linear in weights, identifying my key parameter off of variation in mean county weights.[5] That is,

$$\bar{h}_{ic}(w_j) = \gamma_0 w_j + \gamma_{1i} w_j \cdot \bar{w}_c$$

where $\bar{w}_c$ is the mean weight in county c. In this specification, I assume a fixed parameter on w_j, however, I allow for a random coefficient on the interaction to account for a heterogenous belief structure of vehicle weights. Therefore, a simplified approximation of Equation 2.4 can now be written as:

[4] I define a new vehicle purchase as any vehicle bought in a year less than or equal to that of the model year.

[5] Identification of preference parameters of interest in this paper will not hinge on the non-linear form which county weights ($w \sim f(\cdot)$) enter utility, and thus, will not hinge on identification of $h(\cdot)$. I see a linear specification of $h(\cdot)$ as a reasonable approximation in attaining the expected marginal utilities of interest.

$$
\begin{aligned}
u_{icj} &= -\alpha_i p_j + x_j\beta + \bar{h}_{ic}(w_j) + \varepsilon_{icj} \\
&= -\alpha_i p_j + x_j\beta + \gamma_0 w_j + \gamma_{1i} w_j \cdot \bar{w}_c + \varepsilon_{icj} \\
&= \nu_{icj} + \varepsilon_{icj}
\end{aligned} \tag{2.6}
$$

Until now I have omitted the time dimension for convenience. My data are not merely a cross-section, but a panel. In estimation, I collapse purchase times (observed at the exact day) to a month-of-sample, t. I also allow consumers' preferences for weight to vary based on the mean weights at that point in time ($\bar{w}_{ct}$). As I am only looking at new vehicle purchases, I restrict the choice set at any point in time to that of the model year. That is, a consumer purchases the $j^* \in J_\tau$ that maximizes their utility for that choice set's model year τ.[6] I define a vehicle, j, as a unique make-model in the dataset. Of course there is entry and exit of make-models throughout my time-frame, so J_τ will vary over model years.[7]

Estimation of Equation 2.6 will proceed by Maximum Simulated Likelihood Estimation (MSLE). Purchases are observed at the consumer-level, and thus, estimation will maximize the likelihood of observing those individual purchases. Let the observed purchase of consumer i be noted as y_i. The likelihood of consumer i in county c making purchase y_i is:

$$
s_{icy_i} = \frac{exp(-\alpha_i p_{y_i} + x_{y_i}\beta + \gamma_0 w_{y_i} + \gamma_{1i} w_{y_i} \cdot \bar{w}_c)}{\sum_{k \in J} exp(-\alpha_i p_k + x_k\beta + \gamma_0 w_k + \gamma_{1i} w_k \cdot \bar{w}_c)} \tag{2.7}
$$

I allow consumers to have random coefficients on prices (distributed log-normal) and on the interaction term $w_j \cdot \bar{w}_c$ (distributed normal). That is, for log-normal random variable $\tilde{\alpha}$, with mean parameter $\tilde{\alpha}_0$, and standard deviation parameter $\sigma_{\tilde{\alpha}}$, a random draw from a standard-normal dis-

[6]Whereas I have been notating indirect utility with subscripts icj, utility also varies across dimensions t and τ. Therefore, the correct notation for indirect utility should be written as $u_{icjt\tau}$.

[7]The choice sets in my sample vary in size from 287 to 317, with a total of 529 unique make-models.

tribution, v_r, generates a random draw, α_r; i.e., $\alpha_r = -exp(\tilde{\alpha}_0 + \sigma_{\tilde{\alpha}} v_r)$. Similarly, a random draw for $\tilde{\gamma}_1$ is constructed using an additional draw from the standard-normal distribution, v'_r. That is, $\gamma_{1r} = \gamma_{10} + \sigma_{\gamma_1} v'_r$ where γ_{10} and σ_{γ_1} are the mean and standard deviation of the interaction term's coefficient, respectively. I assume (v_r, v'_r) come from two independent standard-normal distributions, $\phi(\cdot)$. In this manner, I simulate the integral over the vector (v_r, v'_r) from R *iid* draws, such that the likelihood of consumer i in county c making purchase y_i, *unconditional* on random coefficients (α_i, γ_{1i}) becomes:

$$\begin{aligned} s_{icy_i} &= \int \int \frac{exp(-\alpha_i p_{y_i} + x_{y_i}\beta + \gamma_0 w_{y_i} + \gamma_{1i} w_{y_i} \cdot \bar{w}_c)}{\sum_{k \in J} exp(-\alpha_i p_k + x_k \beta + \gamma_0 w_k + \gamma_{1i} w_k \cdot \bar{w}_c)} \phi(v_i)\phi(v'_i) dv_i dv'_i \\ &\approx \frac{1}{R} \sum_{r=1}^{R} \frac{exp(-\alpha_r p_{y_i} + x_{y_i}\beta + \gamma_0 w_{y_i} + \gamma_{1r} w_{y_i} \cdot \bar{w}_c)}{\sum_{k \in J} exp(-\alpha_r p_k + x_k \beta + \gamma_0 w_k + \gamma_{1r} w_k \cdot \bar{w}_c)} \end{aligned} \tag{2.8}$$

I estimate the parameters of the model by maximizing the log-likelihood fuction, over all observations. For random draws of (v_r, v'_r), I use 100 ($R = 100$) Halton draws. I chose $R = 100$ due to the large number of observations, as well as the large choice sets. Halton draws are a popular method for variance reduction, and have been shown provide precise integration with even a small number of draws (e.g., Spanier and Maize, 1991). Bhat (2001) found that 100 Halton draws provided more precise integration than 1000 random draws. Several authors (Train, 2000; Munizaga and Alvarez-Daziano, 2001; Hensher, 2001) have confirmed similar results using different datasets.

My primary specification uses the average transaction price for a particular make-model, j, model year, τ, observed in that county, c (i.e., $p_{j\tau c}$). I include *horsepower* and *price-per-mile*[8] as additional product attributes. *price-per-mile* is used rather than vehicle fuel efficiency, to allow consumer preferences for fuel efficiency to depend on the current gas prices. Including the price of driving in preferences for vehicles will allow me to simulate the arms race feedback loop under different fuel prices. Note that I do not take into account consumer expectations of future vehicle

[8] I compute *price-per-mile* as the current price of gas divided by the vehicle miles-per-gallon. Gas prices are collected from the U.S. Energy Information Administration (EIA), and I use Gulf Coast (PADD 3) monthly prices.

miles traveled (VMT) nor future gas prices (which Anderson et al. (2012) show follow a random walk), therefore my estimated coefficients will contain these expectations, on average. That is, the coefficient will only estimate the *average* marginal utility of *price-per-mile*.

I assume that fuel efficiency is salient, and fully internalized by the consumers, though recent literature suggests that consumers may not fully account for these costs (e.g., Busse et al. (2013); Allcott and Wozny (2014); Sallee et al. (2016); Allcott and Knittel (2018)). In addition, although I am accounting for the price of driving by allowing fuel efficiency to enter utility via gallons-per-mile, it may be worth noting that there is significant evidence that consumers incorrectly perceive driving costs as being linear in miles-per-gallon (Larrick and Soll, 2008; Allcott, 2013). I consider preferences on *price-per-mile* to be a reasonable approximation to true preferences on fuel efficiency, and do not expect this assumption to significantly alter my primary results.

2.6.2.1 Instruments

Directly estimating the parameters in Equation 2.6, ignoring endogeneity of prices and mean weights, could lead to significantly biased estimates. Although, simultaneity should be less of a concern for price endogeneity, since I am estimating based on consumer-level observations, and a single consumer has very negligible influence on the price, it is still possible that I am omitting vehicle characteristics that are correlated with price. That is, ε_{icj} may include unobserved vehicle-level preferences, ξ_j, that are correlated with the price.

Similarly, mean weights in a county $\bar{w}_{ct}$ are endogenous for several reasons. First, since I am estimating preferences that depend on mean weight during the same period of purchase, there are simultaneity concerns. The bias here may not be very significant, however, since thousands of cars make up a county distribution, and it is not likely that one new purchase shifts mean weights by a significant amount. A second and possibly more relevant concern is that of common shocks and autocorrelated common shocks. In particular, I should be concerned with any shocks—which may vary across time and space—that shift preferences for vehicle size for everyone. If these shocks are sustained for periods of time, such as a period of economic growth, I would be concerned that my estimates are biased due to joint shifts in preferences for vehicle size.

To mitigate potential endogeneity concerns, I will use an instrumental variables approach. For price endogeneity, I instrument using the popular instruments of Berry et al. (1995). More specifically, for p_j I use average attributes of all other cars in the same vehicle class (i.e., passenger car, pickup truck, SUV, or van).[9].

Correcting for endogeneity in $\bar{w}_j$ is a bit less trivial. For this, I make use of the fact that fuel-efficiency is a function of vehicle size, and that consumers (by assumption) respond to *price-per-mile* when making their purchase decision. I also exploit a behavioral pattern found in the literature demonstrating that, on average, consumer expectations of future fuel prices follow a random walk (Anderson et al., 2012). Therefore, I can make use of these contemporaneous responses by consumers to gas prices when making their decisions. That is, my instrument will collect all past gasoline prices, contemporaneously, internalized by consumers at their past times of purchase. Under the assumption that consumers account for fuel efficiency and therefore, vehicle size, when purchasing a vehicle, these driving costs will map into weight choices. Thus, by exploiting variation in these past purchase times, and the fuel prices observed, I have plausibly exogenous variation in the weight distribution to which new purchasers respond to.

To illustrate, given preferences specified in Equation 2.6, we know that the expected vehicle size (conditional on observables) purchased by a consumer is $\sum_{j \in J} s_{cj} \cdot w_j$, where s_{cj} are the choice probabilities under the logit formulation in Equation 2.5. Indirect utility depends on current fuel prices, as specified in the *price-per-mile* characteristic, where *price-per-mile* $= f/$*miles-per-gallon*,

[9] That is, for vehicle j in class C, the vector of instruments for p_j is defined as $z_j^p = \frac{1}{|C|-1} \sum_{k \in C \setminus j} x_k$. Here, I include weight in x_j as well, but not the interaction with mean weights, which is endogenous.

and f is the fuel price. Therefore, expected vehicle size is a function of current fuel prices.[10] All consumers on the road made a purchase decision at some previous time, s, and the vehicle size chosen was a function of fuel prices at time s, f_s (Anderson et al., 2012). A consumer's expected vehicle size chosen at time s, conditional on fuel price, is:

$$w(f_s) = E[w|f_s] = E\Big[\sum_{j \in J} s_{cjs} \cdot w_{js}|f_s\Big]$$

Let the set $\mathcal{S}_{ct}$ bet the set of all purchase times for vehicles on the road in county c at time t. Therefore, $s \leq t$ for all $s \in \mathcal{S}_{ct}$. Then the set of all vehicle weights in county c at time t, $\{w\}_{ct}$, is partially a function of past fuel prices internalized at past times of purchase, $\{f_s\}_{s \in \mathcal{S}_{ct}}$. The mean weight in county c at time t, conditional on this set of past fuel prices is:

$$E[\bar{w}_{ct}|\{f_s\}_{s \in \mathcal{S}_{ct}}] = \frac{1}{|\mathcal{S}_{ct}|} \sum_{s \in \mathcal{S}_{ct}} w(f_s)$$

A simplifying, linear specification of $w(\cdot)$ would then make this conditional expectation a function of $\bar{f}_{ct} = \frac{1}{|\mathcal{S}_{ct}|} \sum_{s \in \mathcal{S}_{ct}} f_s$. Using $\bar{f}_{ct}$ as an instrument for $\bar{w}_{jt}$ is attractive in the sense that past fuel prices should not directly impact one's current decision, and we might suspect the timing of these past purchases—which generates weights on this weighted-average of fuel prices—to be orthogonal to unoberved preferences for vehicles. Also note that $\bar{f}_{ct}$ includes information on today's fuel price, f_t, and thus I am assuming this is fully accounted for when including *price-per-mile* in the utility function.

One threat to identification when using $\bar{f}_{ct}$ as an instrument is the possibility that seasonality in fuel prices picks up seasonality in preferences. That is, if we suspect that there is a high degree of seasonality in gas prices, *and* that purchase timings are correlated with preferences for large versus

[10] Increases in f affect consumers' choice of weight. This is because fuel efficiency is partially a function of weight.

small vehicles, we may be concerned, for example, that counties with small vehicles ordinarily tend to purchase in the summer time when gas prices are high. If these patterns exist, then the instrument may simply be picking up county-specific preferences, rather than a causal response to others' choice of vehicle size. To mitigate concerns of seasonality, I exploit large month-to-month changes in fuel prices, as these changes do not seem to be correlated with seasonal trends.[11] The instrument used for $\bar{w}_{ct}$ is:

$$z_{ct}^{\bar{w}} = \frac{1}{|\mathcal{S}_{ct}|} \sum_{s \in \mathcal{S}_{ct}} ln(f_s / f_{s-1})$$

where $ln(\cdot)$ is the natural logarithmic function. The interaction $w_{jt} \times \bar{w}_{ct}$ in the utility function is then instrumented for with $w_{jt} \times z_{ct}^{\bar{w}}$, and the parameters are estimated using the control function approach (Petrin and Train, 2009).

2.6.3 Results

The estimated demand parameters are presented in Table A.6. Columns 1-3 present standard-logit estimates, which do not account for heterogeneity in the coefficients. Column 1 estimates demand assuming no endogeneity in prices or mean weights. Column 2 instruments for prices only, and Column 3 instruments for mean weight and prices. The primary specification from Maximum Simulated Likelihood Estimation is listed in Column 4. As the price coefficients are distributed log-normal, the mean of the distribution is presented in parentheses. These results suggests that accounting for heterogeneity in the parameters does significantly alter the estimates.[12]

[11] I test whether these log differences in the instrument are subject to bias from seasonality by differencing out the fitted values of the instrument when regressed on monthly dummies. The parameter estimates in the standard-logit model are nearly identical.

[12] Note that the identification of the four structural parameters (mean and standard deviations) on the two endogenous variables (price and the interaction of weights) comes not only from my four instruments—three BLP instruments (Berry et al., 1995) and my mean-weight instrument—but also from the different substitution patterns that arise in the data (beyond the standard-logit restrictions on substitution patterns) from variation in the choice sets across model years.

The row labled *WTP(w)/$\bar{w}$* presents the estimated valuation of an arms race, interpreted as the marginal willingness to pay for an additional pound, given a one pound increase in mean weight. These valuations are calculated as $-\gamma_1/\alpha$, and in the case of the random coefficients model, this value is calculated by simulation. These estimates suggest that consumers value a pound \$0.02 more following a one pound shift in mean vehicle size.[13]

To explain the significance of this valuation, consider two very different county vehicle fleets in my dataset—Austin County (a heavy, rural county) and Harris County (a light, urban county). I plot the weight distributions, as of December, 2008, for these two counties in Figure A.6. On average these counties are about 400 pounds different in mean weight. These arms race preferences then suggest that I should value upgrades in size differently between these two counties. Specifically, these estimates suggest I would value an upgrade from a Honda Accord to a Jeep Cherokee about \$3,200 more in the heavy versus the light county.

It is important to note that preferences for vehicle weight are positive within the range of mean weights in the data. That is, because a linear specification is used, the parameter estimate on weight is negative, but these preferences for weight are upward sloping over mean weight. One could specify preferences for weight to enter utility as a non-linear, censored, function of $\bar{w}$, though this would increase the computational burden in estimation. This linear specification only becomes an issue for very small mean weights, which are not observed in the data.

To assess how reasonable my willingness to pay for weight estimates are, I look to Anderson and Auffhammer (2014). Their results suggest that a 1,000 pound increase in *own*-weight in a two car collision, reduces the probability of dying by about 0.1 percentage points. With a 3.65% per year

[13]The instrumental variable estimates of this willingness-to-pay suggest that the initial estimates are downward biased. One possible explanation of this is negatively auto-correlated common shocks in *general* (as opposed to arms race) preferences for weight. For example, if periods of high income growth are followed by lower income growth.

probability of being in a police-reported multivehicle collision (NHTSA, 2007)[14], and the \$7.4 million (\$2006) EPA estimated value of statistical life, this amounts to a per-pound value of about $\frac{0.001}{1{,}000\ lbs.} \times 3.6\% \times \$7.4\ mil. = \$0.27$ per year. At a 5% discount rate, and an average vehicle lifespan of 16 years,[15] this would suggest (in Anderson and Auffhammer's context) a present value of pounds of \$3.2 per pound.

As I have specified preferences for weight to vary based on county mean weight, I can calculate willingness to pay for pounds separately for the heavy and light counties. I calculate value per pound, conditional on mean weight, $\bar{w}$, as $WTP(w|\bar{w}) = \frac{\gamma_0}{-\alpha} + \frac{\gamma_1}{-\alpha} \cdot \bar{w}$. I use the December, 2008 mean weights of the heavy (4,023 pounds) and light county (3,667 pounds). Then using the estimates from Column 4 of Table A.6, I calculate a value per-pound of \$11.35 and \$3.57 for the heavy and light county, respectively. These estimates illustrate quite diverse valuations based on the size of vehicles in the area, suggesting consumers in the heavy county value pounds about three times more than the light county, strictly due to the surrounding vehicles. Considering that Texans generally have a higher preference for large vehicles (and thus have larger mean weights), I see the light county (urban area) as being relatively more representative of the U.S.

2.7 Counterfactuals

In this section I focus on isolating the arms race effect through a series of counterfactual experiments. In each experiment, I assign simulated consumers a sustained period of high or low gas prices, proxying for static gas taxes. I will then examine the long-run effects of an arms race in each scenario. The purpose of this exercise is to examine the extent to which dynamic preferences for vehicle size matter when imposing a static Pigouvian tax on vehicles.

To isolate the impact of an arms race, I will simulate purchases under two different utility specifications—

[14]This probability of being involved in an accident was calculated by Anderson and Auffhammer (2014) as the total number of registered vehicles in 2005 divided by the total number of vehicles involved in a reported multivehicle collision.

[15]The EPA assumes the average vehicle travels 15,000 miles a year. A Forbes article (Gorzelany; March, 2014) states that cars are now built to last up to 250,000 miles.

an arms race group and a no arms race group. In each group, I simulate purchases in every period (one month) assuming a constant turnover rate. To illustrate, suppose consumers in a representative county at time $t = 0$ make up a vector of weights, W_0, belonging to that county's weight distribution. Given preferences for weight in an arms race, vehicle purchases at $t = 0$ depend on the mean of this vector, $\bar{w}_0$. The simulation then proceeds by replacing a subset of W_0 with the vehicle sizes produced from these simulated purchases.

At $t = 1$, the arms race group and the no arms race group respond differently. A portion of the population in the arms race group now observes the new weight vector, W_1, and makes their purchase decision based on the new mean weight, $\bar{w}_1$. However, the no arms race group does not observe the newly updated vector. They are essentially "blindfolded" from observing the updated weights around them. For all periods, the no arms race group continues responding to $\bar{w}_0$, which corresponds with static preferences for weight. The feedback loop then continues in this manner for the two groups up to a pre-specified period T.

The indirect utility functions for these two groups can be written as follows for the arms race and no arms race group, respectively:

$$\begin{aligned} u_{icjt}^{AR} &= \alpha_i p_{jt} + x_{jt}\beta + \gamma_0 w_{jt} + \gamma_{1i} w_{jt} \cdot \bar{w}_{ct} + \varepsilon_{icjt} \\ u_{icjt}^{\neg AR} &= \alpha_i p_{jt} + x_{jt}\beta + \gamma_0 w_{jt} + \gamma_{1i} w_{jt} \cdot \bar{w}_{c0} + \varepsilon_{icjt} \end{aligned} \tag{2.9}$$

The primary outcome will examine the county average weights generated over time by the two groups. I will examine the difference between the two groups under different fuel prices (affecting preferences through *price-per-mile*), in anticipation that this "arms race effect" will be exacerbated under very low gas prices, and can be internalized under higher gas prices.

The simulations will be examined under two different, empirical, initial weight distributions. I begin with the weight distributions as observed in December, 2008. At this point in time, gas prices were very low (\$1.665). The two distributions I will use come from Harris (urban) and Austin (rural) County, Texas—two counties in the tails of the weight distribution (see Figure A.6). My representative choice set, which I will hold constant over periods simulated, will be that of model year 2009.

These distributions will be the basis of the initial weight vectors, W_{H0} and W_{L0}, for the heavy and light counties, respectively. Rather than using the entire vectors of weights (which are quite large), I draw $N = 10,000$ representative weights from each distribution to make up these weight vectors. I use a turnover rate of 2%—consistent with the observed household purchase rates.

Figure A.7 presents the simulation results for the two counties under sustained periods of low gas prices ($T = 144$, months). These graphs illustrate the mean county weights for the arms race and no arms race groups, isolating the feedback loop. The difference in outcomes between the two counties come strictly from the difference in initial starting weights. Recall that the no arms race group continues to respond to mean weights at $t = 0$, and $\bar{w}_{H0} > \bar{w}_{L0}$ implying a higher static preference for weight in the heavy county.

The difference in outcomes between these two groups is significant. The heavy county at time T has a mean county weight of 436 pounds more for the arms race over the no arms race group (about the difference between a large sedan and mid-sized SUV). With very low gas prices, even the no arms race group increases its vehicle sizes. In the arms race group, consumers not only respond to the lower fuel prices, but they respond to the previous periods' consumers' choices of vehicle size.

The light county produces a much smaller response, increasing their average weight by 78 pounds over the no arms race group. This county starts out with a lower initial weight, so consumers do not see a large benefit in increasing their weight relative to others. In the arms race mechanism, the consumers have little incentive to protect themselves from the relatively small sized vehicles around them.

Figure A.8 presents results from a similar simulation, except now under sustained periods of high fuel prices. I set the price of gas at $4 a gallon and forward simulate purchases in the same manner. In this case, there is a significantly higher marginal cost of increasing one's vehicle size.

In the case of high gas prices, it is now more costly for consumers to increase their vehicle size. They must weigh the option of strategically responding to other large vehicles on the road, versus the higher cost of driving a large car. In the heavy county, initial weights are sufficiently high to still generate an arms race effect, though a much smaller one. By time T, county average weights are now 193 pounds heavier in the arms race group than the no arms race group; not quite a full shift in vehicle class.

The simulation for the light county produces an interesting outcome of a "reverse arms race". By time T, consumers in the arms race group now have an average weight of 96 pounds *less* than that of the no arms race group. This may seem counter-intuitive, as it seems that consumers in an arms race should always want larger vehicles than their neighbors. To illustrate the intuition of this outcome, consider that in period $t = 0$—when gas prices are sufficiently high and initial weights are sufficiently low—the marginal benefit of an additional pound, via the influence of others' weights, does not overcome the marginal cost, via the high fuel prices. Consumers in this case are not willing to pay to increase their vehicle size. At $t = 1$, consumers in the arms race group observe the high cost of increasing vehicle size, and also observe the decrease in mean weights from $t = 0$ small vehicle purchases. The arms race mechanism suggests that these consumers are less concerned with their protection from the less threatening, smaller vehicles on the road. Given the high fuel prices, these consumers are now even less willing to increase their weight, so they are better off decreasing their weight even further. This type of behavior continues in future periods, as vehicle weights spiral downwards.

This exercise verifies that with sufficiently high gas prices, an arms race can be reversed. This result has tremendous policy implications, as any objective aimed at reducing the externalities produced by large vehicles may be able to exploit this feedback loop. Note, however, that I have ignored the role of endogenous prices in these simulations. Sustained periods of high or low gas prices should impact the prices of vehicles (e.g., Busse et al., 2016), and so should changes in consumers' value of weight. In Online Appendix B.7, I allow firms to adjust their prices to this change in consumers' value of vehicle size by assuming that firms participate in static Bertrand price competition in every period. Simulations show quantitatively similar results when consumers observe new Bertrand-Nash equilibrium prices each period.

In Online Appendix B.6, I compute optimal, period-specific, gasoline taxes in order to fully internalize an arms race. These taxes will be dynamic, in order to adapt to the dynamic preferences for weight in an arms race. The key insight from this exercise demonstrates how policy can exploit an arms race. The results in Figure A.8, Panel B, indicate that policy can strategically manipulate arms race preferences to achieve tremendous environmental benefits. Optimal tax estimates, which incentivize the arms race group to consume the same vehicle sizes as the no arms race group, directly show this.

Figure B.6, Panel A, in Online Appendix B.6, suggests that a tax must initially start out high, to induce smaller car purchases in a low gas price environment. Placing a higher tax on gas increases the cost of driving large cars, thus internalizing arms race incentives. When size distributions are sufficiently small, this eventually kick-starts a "reverse arms race". Therefore, the optimal gas tax can ultimately be reduced, and in the case of the light county in Figure B.6, the tax may ultimately be eliminated.

2.8 Conclusion

There is an obvious relationship between preferences for safety and preferences for size, with respect to vehicle demand. Measuring the extent to which consumers favor these attributes is essential in evaluating environmental policy. As larger vehicles on the road create higher levels of carbon emissions and an increased risk to others, policy efforts should focus on the incentives that drive the adoption of these vehicles. I examine an apparent motivating factor in the form of a vehicle arms race, identified through salient outcomes of large vehicles. In addition, I study how consumers' preferences for vehicle sizes change as a function of the cars around them. These two components of an arms race describe how preferences for weight are not static, but dynamic, providing significant implications for Pigouvian policy.

This paper has presented strong evidence that households respond to salient information regarding vehicle safety. This type of information was presented as a response to fatal accidents harming a next-door neighbor. The consumer response identified was a heavier household vehicle fleet with lower fuel economy. These responses were most likely made with the purpose of protection from opposing drivers; but this type of behavior produces very costly externalities.

Demonstrating that the consumers respond to a salient shock to their preferences for safety by purchasing larger vehicles suggests that there are dynamics at play with respect to these preferences. If everyone else on the road is driving Hummers, it is not conducive to purchase a Mini. Consumers with a strong preference for their own protection take this information into account, and best respond to other vehicles on the road.

I demonstrate how consumer preferences for vehicle size vary based on the size of vehicles on the road by estimating a discrete choice model of vehicle purchases, and allowing for county weight-specific preferences for size. Estimating this model is useful for two main purposes. First, the parameters of the model allow me to place a dollar value on relative weight. Specifically, my estimates suggest that a household in a Texas county with a high average vehicle size values an upgrade from a Honda Accord (Sedan) to a Jeep Cherokee (SUV) approximately \$3,200 more than they would value the upgrade in a Texas county with a low average vehicle size.

The parameter estimates of this model also allow me to simulate the long term effects of an arms race under different gasoline prices. These exercises demonstrate that an arms race may exacerbate vehicle size increases in a low gas price environment, but an arms race may be reversed with a sufficiently high price of gas.

The findings in this paper suggest that dynamics matter in the context of Pigouvian taxes. Not accounting for an arms race could lead to a significant underestimate of an optimal tax. Therefore, understanding the extent to which preferences for vehicle size are dynamic, as in an arms race, is a crucial element in evaluating Pigouvian taxes.

3. MOVING TO FLOOD PLAINS: THE UNINTENDED CONSEQUENCES OF THE NATIONAL FLOOD INSURANCE PROGRAM

3.1 Introduction

Since the creation of the National Flood Insurance Program (NFIP), the U.S. government has paid out over $51 billion to cover flood losses. Almost half of these payouts went to just 25 counties, which are also among the fastest growing counties by population (Kane and Puentes, 2015). There are a number of potential explanations for this. The aesthetic appeal of coastal living may have encouraged households to increasingly move to and stay in coastal counties, which are likely to be flood-prone (Kahn, 2005; Boustan, Kahn, and Rhode, 2012). Because most of US economic activity is concentrated on its ocean and Great Lakes coasts (Rappaport and Sachs, 2003), there may be labor market incentives to locating in these areas. We focus on a previously unstudied factor: that insuring people against potential flood losses contributes directly to population growth in flood-prone areas.

The motivation to provide insurance against the consequences of flooding is clear: Globally, the costs of weather-related natural disasters is increasing over time, from $8.9 billion during the early 1980s to $45.1 billion in more recent years (Bouwer et al., 2017). Nationally, severe weather-related disasters appear to be linked with increased out-migration, poverty, and lower home prices (Boustan, Kahn, Rhode, and Yanguas, 2017). Moreover, the economic effects of these disasters persist years afterward; Strobl (2011) finds that counties affected by hurricanes experience a significant reduction in their annual economic growth rate. In response, the federal government has offered significant financial assistance to victims of flooding. Deryugina (2017) documents that direct disaster aid provided to affected counties amounts to $155-$160 per capita, and that they also receive a further $780-$1,150 per capita through non-disaster social insurance programs in the ten years following a hurricane.

Given the amount of federal aid targeted toward flood victims, a natural question to ask is whether this insurance has induced people to move to or stay in flood-prone areas. This moral hazard response would increase the burden on taxpayers, adding to the program's already existing inefficiencies (Kahn and Smith, 2017). Though important, it is difficult to identify the causal pathway as joining the NFIP is voluntary. Since many factors may make communities more or less likely to join, using actual NFIP adoption to define treatment may lead to biased estimates due to reverse causality. For example, increased migration into an area could alter its flood risk or increase its relative demand for coverage, making it more likely to enroll into the program.

In this paper, we instead test for potential moral hazard from the NFIP by exploiting several programs that the Federal Emergency Management Agency (FEMA) implemented to pressure risky communities to join the NFIP. FEMA is mandated to identify and map out areas with elevated flood risks. It then sanctions areas identified to have elevated flood risk if they do not join the NFIP within a year of their flood map being published. We examine how population flows change in more risky counties relative to less risky counties, using the publication timing of the flood maps published by FEMA to determine treatment timing. Because our specifications account for state-by-year fixed effects, our identifying assumption is that absent the pressure from FEMA to join the NFIP, these flood-prone counties would have experienced population changes similar to what other counties in the same state experienced. We present evidence in favor of this assumption by demonstrating that more risky and less risky counties follow similar trajectories in migration patterns prior to the publication of flood maps, and only start diverging after.

Results indicate that flood insurance availability affected the decision to live and stay in more flood-prone areas. We estimate that population in counties whose communities are incentivized to join the NFIP increase by 4 to 5 percent. This effect is primarily driven by existing residents choosing to remain in these high-risk areas when they receive federal flood protection. By comparison, we find little (but some suggestive) evidence of increases in migration to flood-prone areas. This is consistent with existing residents being more informed about the availability and need for federal flood insurance. These estimates are robust to various specifications, including adding time-varying controls for county-level income and employment characteristics, county-specific linear time trends, and to accounting for water-related natural disaster events which might induce take-up.

We provide further evidence of moral hazard by estimating the heterogeneous effects of NFIP by underlying flood propensity. To do so, we use cross-sectional information on historical flood risk to examine the treatment intensities of NFIP across different levels of empirical risks. Whereas our original estimates measure a treatment effect on the primarily risky counties in which FEMA directly targeted, our heterogenous estimates incorporate actual empirical risk levels into these treatment effects. Using this approach, we estimate that NFIP produces an additional 4.4 percent increase in population for a one standard deviation increase in flood risk. As before, we attribute most of this effect to current residents choosing to stay in, rather than move out of, more flood-prone areas; though we do find some evidence of increased migration into these areas.

This study contributes to the literature in the following ways. First, to our knowledge, we provide the first causal evidence of the effect of flood insurance availability on population growth in flood-prone areas. In doing so, we complement existing work on the take-up and effects of flood insurance, which document a positive relationship between flood risk and flood insurance demand in coastal counties (Landry and Jahan-Parvar, 2011; Kriesel and Landry, 2004). Gallagher (2014), and Kousky and Shabman (2014) find that the flood insurance take-up increases in the year following hurricanes and large flooding events, even though the underlying flood risk in the area did not change. Relatedly, Gregory (2014) estimates a dynamic discrete choice model to show that the Louisiana Road Home grant program increased the rebuilding rate in New Orleans after Katrina.

More broadly, because our results suggest that people take on more risk, we also contribute to the larger set of research on the moral hazard responses to insurance availability. Such responses have been found to occur with health insurance (e.g., Einav, Finkelstein, Ryan, Schrimpf, and Cullen, 2013; Spenkuch, 2012; Keane and Stavrunova, 2016), life insurance (Cawley and Philipson, 1999), and automobile insurance (Weisburd, 2015; Dionne, Michaud, and Dahchour, 2013).

Finding evidence of a moral hazard response to flood insurance availability has important policy implications. At present, the NFIP covers over \$1 trillion worth of property (Michel-Kerjan, 2010). Total damages have increased together with incurred losses over the years, raising concerns over the financial viability of providing subsidized flood insurance. Assuming flood damages are proportional to population size, our estimates suggest that the moral hazard produced by the NFIP has been responsible for significant costs from major historical floods, such as those coming from Hurricane Katrina. To illustrate the size of moral hazard costs, we examine recent flooding in New Orleans Parish, Louisiana, which ranks in the 75th percentile in historical flood risk, and Harris County (Houston), Texas, which is in the 90th percentile. The NFIP spent over \$16 billion in insurance payouts as a result of Hurricane Katrina, much of which went to New Orleans, a city partially below sea level (Michel-Kerjan and Kousky, 2010). Our triple-differences estimates suggest that the moral hazard produced by the NFIP led to costs that were 6.6% higher than they would otherwise have been. As for Harris County (Houston), we estimate that the NFIP was responsible for approximately 15% of the damages from Hurricane Harvey.

In addition to the increased costs incurred from past major disasters, the perverse incentives created by the NFIP play a major role in inhibiting adaptation to the future risks of climate change (Barreca et al., 2016). We show that NFIP adoption is a strong driver of population growth in high-flood risk areas, adding to the already growing costs of increasingly frequent climate change driven natural disasters. Our finding that there are large moral hazard costs of insuring populations against risks of floods suggests large inefficiencies in the rate structure of the program. Therefore, going forward, policy must aim to better incorporate heterogeneous flood risks into insurance premiums.

The remainder of the paper proceeds as follows: Section 3.2 discusses the relevant background of the National Flood Insurance Pogram. In Section 3.3, we present our empirical strategy to estimate the magnitude of moral-hazard from flood insurance. In Section 3.4, we discuss the data used in this paper. In Section 3.5, we present our primary results. In Section 3.6, we extend our primary estimates to allow heterogeneous effects by risk level. Finally, in Section 3.7, we conclude with a brief discussion of our findings.

3.2 Background

The National Flood Insurance Program (NFIP) was created through the National Flood Insurance Act of 1968. Before the NFIP, private insurers were largely unable to offer flood insurance, both because the necessary flood risk maps did not exist and because actuarially fair premiums were thought to be too expensive for prospective buyers (Anderson, 1974). Beginning in 1973, various federal programs were implemented to systematically identify and price underlying flood risks. The Flood Insurance Administration in the Department of Housing and Urban Development created the earliest flood maps, which were called Flood Hazard Boundary Maps (FHBMs). While these maps are not as detailed as present-day flood maps, they indicate which communities were more flood-prone before flood insurance was available. Most of these flood-prone communities were identified between 1973 and 1978.[1]

In 1979, the responsibility of creating flood maps was transferred to the Federal Emergency Management Agency (FEMA). FEMA creates Flood Insurance Rate Maps (FIRMs) rather than FHBMs. Relative to FHBMs, FIRMs depict more levels of flood risk within mapped communities (Morrissey, 2006). The level of detail on FIRMs allows premiums to vary by the riskiness of the zone in which the property is located, and identifies which properties are required to carry flood insur-

[1]While a few (less than 3 percent) FHBMs were created before 1978, our understanding is that these were only created for emergency or temporary enrollments into the NFIP following a major disaster. Similarly, some maps were created before 1973. (Morrissey, 2006).

ance. FEMA is also mandated to evaluate the need to update or create more flood maps at least once every five years. Thus, while roughly half the communities in our sample were mapped or re-mapped during the first ten years of FEMA, some communities received their first FIRMs after 1989. The last big push to create and update flood maps relevant to our study period was in 1997, when FEMA started the Flood Map Modernization Initiative. The goal of this program was to transition from paper maps to digital flood maps, and to create new digital flood maps where necessary. [2]

The flood risks depicted on maps are based on the outcome of flood hazard studies. These studies use geophysical and environmental data, land and aerial surveys, and interviews with the local population. Importantly, the flood maps we use in our paper are based on historical data and by law could not be based on future flood projections, or factors that affect future flood risk such as expected population growth and development (Pralle (2017); TMAC (2015)). In fact, as late as 2015, a federal advisory committee recommended that “FEMA should use population growth as an indicator of areas with increased potential flood risk."[3] As we use data prior to 2011, this means that communities that have higher expected future flood risk, whether due to anticipated construction or population growth, should not be more likely to receive flood maps.

Several flood disaster-related policies have made flood map publication a strong driver of future NFIP enrollment. The Flood Disaster Protection Act of 1973 requires that communities join the NFIP within one year of being identified by FEMA as flood-prone, or be sanctioned. These sanctions make the affected communities ineligible for most forms of disaster assistance, and prevent property owners in those communities from acquiring flood insurance. Communities who do not have flood hazard areas identified by FEMA can still voluntarily join the NFIP, but staying out of

[2] Some FIRMs were created before 1979, but these are also just hand-drafted emergency maps made after a major flood (Morrissey, 2006).

[3] The Technical Mapping Advisory Council recommended that FEMA incorporate future flood projections into mapping in late 2015, meaning any changes would have been made in 2016 or later. Our analysis uses only data up to 2011, before FEMA would have incorporated factors that affect future flooding, such as population growth and construction, in measuring flood risk.

the program does not open them to the same sanctions (Michel-Kerjan, 2010). Property owners in special flood hazard areas (or 100-year floodplains) who have federally-backed mortgages are mandated to obtain flood insurance, conditional on their communities joining the NFIP. However, because this policy was not strictly enforced through the 1980s, take-up of flood insurance within NFIP communities was low.

In the 1990s, two policy changes occurred that made map publication and updates a stronger predictor of future NFIP enrollment. In 1994, the Riegel Community Development Regulatory Improvement Act was passed. This new law penalized mortgage lenders for not ensuring that borrowers had flood insurance whenever it is required. Second, FEMA conducted Cover America, which was an extensive information campaign from 1994 to 2000. This resulted in greater awareness about the flood risks and sanctions from not joining the NFIP (Chivers and Flores, 2002). Figure A.9 illustrates how take-up increased from the 1990s onward. Today, the NFIP covers about $1.2 trillion worth of property (Michel-Kerjan and Kunreuther, 2011). The coverage limits are $250,000 for residential buildings and $500,000 for non-residential buildings (Burby, 2001).

3.3 Empirical Strategy

This paper examines the moral hazard created by the NFIP. In this context, we define moral hazard to occur whenever households participate in more risky behavior—for example, moving to and staying in risky areas—when insured against the potential costs of this behavior. As FEMA primarily targets high risk areas to join the NFIP, finding evidence of moral hazard hinges on correctly identifying the treatment effect of NFIP on the treated areas. However, an approach that directly uses NFIP adoption to construct the treatment variable may produce biased results due to potential reverse causality. That is, as we are primarily interested in the causal effect of NFIP enrollment on population growth, any underlying causal effects of population growth on NFIP enrollment would bias our estimates.

To overcome this problem, we exploit FEMA's direct targeting of risky areas following the Flood

Disaster Protection Act of 1973. We use the Flood Hazard Boundary Map (FHBM) assignments in the 1970s to identify treatment. This group of targeted communities strongly correlates with the communities that ultimately enrolled into the program, as well as empirical flood risk across areas (see Figure A.12). Years (and often decades) later, FEMA followed up with largely the same group of communities by upgrading them to Flood Insurance Rate Maps (FIRMs), which describe the rate structure communities would face if they enroll into the program. In our sample, 99% of counties with at least one FHBM ultimately received a FIRM, providing evidence that this targeted group remained consistent.

Since these communities were given one year to join the NFIP after receiving a FIRM before being sanctioned, we exploit this as a plausibly exogenous incentive which induced many of the communities to enroll into the NFIP. This strategy should be robust to anticipated changes in population, as the timing of initial FIRM assignment was independent of these expectations (as described in Section 3.2). We demonstrate the relevance of this proposed instrument by estimating the following first stage equation:

$$postNFIP_{cst} = \alpha \cdot \textit{postFIRM-FHBM}_{cst} + x_{cst}\tilde{\beta} + \tilde{\lambda}_{st} + \tilde{\gamma}_{cs} + \tilde{\varepsilon}_{cst} \tag{3.1}$$

where $postNFIP_{cst}$ indicates actual enrollment into NFIP for county, c, in state, s, at year, t. This variable represents the proportion of a county which has been enrolled, and is thus between zero and one. *postFIRM-FHBM* describes the proportion of a county assigned an FHBM that has also been assigned a FIRM. α estimates the relationship between FEMA targeting and actual take-up. x_{cst} is a set of time-varying, county-level controls, including information on employment, income, and natural disasters. We include a set of state-by-year and county fixed effects, λ_{st} and γ_{cs}, respectively.

In the reduced form, we implement a difference-in-differences approach and compare high-risk areas—proxied by FHBM counties—to low-risk areas, before and after initial FIRM assignment. As FHBM communities do not encompass the entire population of NFIP enrollees, our approach also makes use of an instrumental variable design. Our identification requires that changes in population over time in the FHBM counties would track closely with non-FHBM counties, absent FIRM assignment, and that FEMA interventions affect our primary outcomes, only through NFIP enrollment. Our primary estimating equation for the intent-to-treat (ITT) of NFIP on population flows is the following:[4]

$$migration_{cst} = \delta \cdot \textbf{\textit{postFIRM-FHBM}}_{cst} + x_{cst}\beta + \lambda_{st} + \gamma_{cs} + \varepsilon_{cst} \tag{3.2}$$

Our main outcome is the natural logarithm of population. We also decompose the estimated changes in population into the log- number of non-migrants each year, as well as the log- number of inflow migrants. δ is our coefficient of interest, estimating the ITT effect of NFIP on the population outcomes. Causal interpretations for NFIP enrollment should be made on the scaled coefficient, δ/α.

Because we use FHBM assignment to proxy for risky counties, we interpret a positive effect on population ($\delta > 0$) from these areas being pressured to join the NFIP as evidence of moral hazard. That is, if we observe that population in risky counties increases after the publication of their communities' flood maps, our interpretation is that impending flood insurance availability caused individuals to believe that they could be compensated for future potential flood losses by the government, and therefore became more likely to take on the added risk of moving to and staying in flood-prone areas.

[4]Our specification defines our "treatment" group based on FHBM assignments in the 1970s. These group assignments map closely to FIRM assignments, decades later. Specifically, 99% of counties with at least one FHBM have at least one FIRM ultimately assigned to them.

In addition to our main difference-in-differences specification in Equation 3.2, in Section 3.6 we also allow for heterogeneous effects within treatment by implementing a triple-differences (triple-diff) estimator. In Figure A.12, we document a relationship between treatment assignment and empirical risk. Because of this, our primary strategy should by itself suggest evidence of moral hazard. However, we provide further evidence by exploiting additional variation in the treatment group's empirical risk level. To do this, we extend Equation 3.2 by allowing the treatment effect of NFIP to vary based on a measure of actual risk—county-level historical flood propensity. Formally, we estimate the following equation:

$$\begin{aligned} migration_{cst} = &\delta_0 \cdot \textit{postFIRM-FHBM}_{cst} \\ &+ \delta_1 \cdot \textit{postFIRM-FHBM}_{cst} \times \textit{flood risk}_{cs} \\ &+ x_{cst}\beta + \mu_t \cdot \textit{flood risk}_{cs} + \lambda_{st} + \gamma_{cs} + \varepsilon_{cst} \end{aligned} \tag{3.3}$$

where *flood risk*$_{cs}$ is our measure of average annual flood episodes for county-state, cs. We control for time-varying confounders specific to flood-prone areas, $\mu_t \cdot \textit{flood risk}_{cs}$. All other pair-wise interaction terms are absorbed into fixed effects. δ_0 measures the constant ITT effect of NFIP when *flood risk* is equal to zero.

In this specification, δ_1 is our coefficient of interest. It measures the additional ITT effect of NFIP coming from one additional flood episode per year. In this sense, δ_1 estimates the extent to which NFIP produces moral hazard. An estimate of $\delta_1 > 0$ implies the NFIP produces a larger increase in population in historically risky areas, suggesting that households internalize this reduction in risk through coverage of potential losses. We define a county's flood risk according to the average annual flood episodes experienced in that area over the time-span of our data, as reported by NOAA.[5]

[5]Due to data constraints, we use in-sample floods instead of pre-NFIP floods. We discuss and test for potential endogeneity in Online Appendix B.9.

3.4 Data

3.4.1 Population

To implement our empirical strategy, we require data on population and migration over a long time period. We obtain this from county-to-county migration files published by the Internal Revenue Service for years 1990 through 2011.[6] These files are constructed from individual tax returns each year. By tracking changes in addresses, the IRS is able to track the number of people making inter-county moves between two filing years as well as the number of people that stay in their county. Because tax returns are filed every year, the IRS data are arguably the best source of data on movers over a long time period. As is standard, we construct all of our population outcomes using the number of exemptions to proxy for the number of people. From these files we construct county-level population, which is the sum of non-migrants (residents that did not change counties) and inflow migrants (new residents who moved from another county). Importantly, although our data do not cover the first two decades of the NFIP, they do cover the entire period when flood map publication became a stronger predictor of NFIP adoption (see Figure A.9).

These data have some potential limitations, which we account for in this paper. First, for confidentiality the IRS does not report totals based on fewer than 10 tax returns. While this does not affect our main results, it prevents us from conducting additional analysis, such as examining where in-migrants are coming from, or where out-migrants are moving to. Second, there were methodological changes in 2011 that led to an increase in the number of tax returns that were being counted in the county-to-county migration files. Because the resulting increase in tax returns was not allocated uniformly across counties (Pierce, 2015), we exclude the entire affected period from our main results. Finally, these data will not reflect moves by those individuals not required to file an income tax return.

[6]Data are available beyond 2011, but methodological changes make it inappropriate to link data post-2011.

3.4.2 National Flood Insurance Program

Constructing our instrument requires that we know which communities were identified as risky at the start of the NFIP, and when they were pressured by FEMA to join. We obtain this information from the Community Status Book published by FEMA. The data contain information on the publication dates of community-level Flood Hazard Boundary Maps (FHBM) and Flood Insurance Rate Maps (FIRM), and the dates that communities adopted the NFIP or were sanctioned by FEMA for not joining the NFIP. The inclusion of the publication date of the FHBM is important because it allows us to identify the communities that were initially identified during the watershed analyses of the 1970s as having elevated flood risk. Likewise, the publication date of the FIRM tells us when the community was next pressured by FEMA to join the NFIP.

We aggregate community-level information on map coverage to the county-level by constructing a variable for the fraction of a county's communities that have a FHBM.[7] An issue which sometimes arises when aggregating up to county is that communities and counties can overlap. That is, counties may contain multiple communities, and communities may contain multiple counties. For simplicity, we treat each community as identical and aggregate communities within counties. In doing so, it is possible that one community appears in multiple counties. As this will most likely show up as classical measurement error in an explanatory variable, it may attenuate our point estimates for effects at the county-level.

Aggregating from communities up to counties also implies that we will have fractional treatments at the county-level. This means that our treatment indicator, *FHBM*, will be a value between 0 and 1, indicating the fraction of communities within a given county with a FHBM. Similarly, our *post-FIRM* indicator will describe the fraction of a county that is in a period following FIRM assignment.

[7]FEMA defines relevant areas as "communities," and NFIP enrollment occurs at this level. Therefore, we maintain this terminology in this paper; however, these "communities" are simply towns and cities.

Figure A.10 presents the distribution of counties by the fraction of their communities with a FHBM. More than 70 percent of counties have at least one community under FHBM, and 45 percent of counties are comprised of entirely FHBM communities. However, this does not imply that the majority of counties are risky, since typically, only a small fraction of land area covered by a FHBM is identified to be a Special Flood Hazard Area. According to the Federal Emergency Management Agency (1983), only 4% of the total U.S. land area is within the 100-year floodplain (Maantay and Maroko, 2009; Robinson, 2004). Nevertheless, a county's position in the distribution is indicative of its assessed flood risk relative to other counties.

As our data are aggregated to the county level, we interpret the fraction of communities with an FHBM as a measure of flood risk. That is to say, the larger the fraction of a county with an FHBM, the higher the risk of flood. For example, 65% of communities in Salt Lake City county have FHBM. This places the county in the bottom quartile of counties by fraction with FHBM, which corresponds with its low empirical flood risk. By comparison, all communities in Orleans Parish (New Orleans), a very flood-prone county, have FHBM. The relationship between FHBM and empirical flood risk can be seen in Figure A.12. This plots our county FHBM measure against a cross-section of mean-annual floods, as reported by the National Oceanic and Atmospheric Administration (NOAA). We show that there is a positive and statistically significant relationship between these two measures of flood risk.

3.4.3 Other Data

We also require a set of time-varying county-level controls in order to test the robustness of our estimates. We use county-level data on natural disasters from the FEMA Disaster Declarations Summary. The data contain a list of counties for which the state governor requested and was granted a federal disaster declaration for natural disaster. Natural disasters include fire and water-related events such as hurricanes. A disaster declaration allows counties to receive disaster assistance.

Over 75% of flood-related requests are approved, and the request for a federal disaster declaration requires documentation of damage assessments. The lack of information on declined requests for disaster declarations is a potential limitation of using this data. However, this data is still considered the most comprehensive source of large flooding events.[8] In addition, we use annual flood episodes, reported by National Oceanic and Atomospheric Administration (NOAA) for our analysis in Section 3.6. Finally, we obtained county-level income and employment data from the Bureau of Economic Analysis.[9]

3.5 Results

3.5.1 First Stage Estimates

We begin our analysis by testing the relevance of our instrument for NFIP enrollment. As FHBM communities were strongly incentivized to join the NFIP following FEMA's FIRM assignments, we anticipate a near one-to-one relationship. It is important to note however, that communities not directly targeted can enroll into the program as well. Because of potential reverse causality, this is one particular source of variation which we are not identifying off of.

Estimates for Equation 3.1 can be seen in Table A.8. Standard errors are clustered by county. The estimates are highly significant, producing an F-statistic exceeding 100. This result assures us that our instrument is relevant. The estimates are also not sensitive to the inclusion of additional controls, where estimates do not change at the thousandths decimal place after controlling for demographic variables (Column 2), or declared disaster controls (Column 3).

The first stage estimates suggest that 95 percent of counties targeted by FEMA enroll into the NFIP. As our first stage estimates are close to 1, we will interpret our reduced form estimates (i.e., of Equations 3.2 and 3.3) as if they were estimates from the structural equations, though the scaled

[8] See Gallagher, 2014 for a discussion of other data on flooding events.

[9] The BEA also publishes intercensal estimates of county-level population. However, the nature of the data does not allow us to account for dynamic effects of the NFIP in the way that the annual IRS data does.

estimates of the effect of NFIP will be slightly larger.

The dynamics of the first stage can be seen in Figure A.13, where we regress NFIP on a series of lagged and leading terms of our instrument. The final lagged term represents the effect through the rest of the data so that the estimates are relative to the pre-intervention periods. These estimates illustrate the strong incentives that FEMA imposed on affected communities, inducing the majority of take-up in the first year following the intervention.

3.5.2 Graphical Evidence

Before presenting our main results, we present evidence to support the parallel trajectories assumption of our difference-in-differences approach. We do this by estimating a fully dynamic version of Equation 3.2, with several lagged and leading terms for FIRM assignment. Specifically, we estimate the following equation:

$$\begin{aligned} migration_{cst} = &\sum_{l=-\underline{L}}^{\bar{L}-1} \delta_l \cdot \textit{newFIRM-FHBM}_{cst-l} + \delta_{\bar{L}} \cdot \textit{postFIRM-FHBM}_{cst-\bar{L}} \\ &+ x_{cst}\beta + \lambda_{st} + \gamma_{cs} + \varepsilon_{cst} \end{aligned} \tag{3.4}$$

where $\bar{L}$ is the number of lags and $\underline{L}$ is the number of leads. This equation estimates the dynamic effects at each point of take-up for FHBM counties, *newFIRM-FHBM*, which indicates the fraction of communities in a county enrolling at a given point in time. These coefficients are estimated relative to the pre- period, as the final lag, *post-FIRM*, estimates the average effect through the end of the data. The leading terms serve as placebos, as we should not expect to see responses to future period treatments, and thus, formally tests our parallel trends assumption.

To illustrate the endogeneity problem in using actual NFIP take-up, we implement a difference-in-differences strategy using a naïve version of Equation 3.4.[10] Using actual NFIP take-up instead of our instrument also allows us to directly test for reverse causality in a *Granger*-causality sense. The naïve estimates are presented in Figure A.14. For all outcomes, there is an obvious divergence in treatment prior to NFIP enrollment. For effects on population (Panel a) and non-migration (Panel b), these estimates suggest migration changes up to three years prior to enrollment may have contributed to future NFIP take-up.

With sufficient evidence that invalidate the naïve approach, we directly test for reverse *Granger*-causality in our primary design, instrumenting for NFIP with FEMA interventions. Estimates from Equation 3.4 are presented in Figure A.15. It is immediately clear that for population (Panel A) and non-migration (Panel B) outcomes, exploiting FEMA map publication timing allows us to circumvent the same biasing factors that partially determine NFIP adoption. On the other hand, Panel C shows that our in-migration outcome still exhibits suggestive evidence of positive, though insignificant, divergence in the year prior to treatment. These inflow data make up a much smaller proportion of total population levels than our non-migrant outcome, and thus, may be more susceptible to noise. For these reasons, we will primarily focus on our *population* and *non-migration* outcomes. Furthermore, in Section 3.5.3, we will show that in many specifications, static estimates of the effect of NFIP on *in-migration* are not statistically significant.

3.5.3 The Causal Effect of NFIP on Population Flows

Table A.9 presents the reduced form estimates from our Equation 3.2, our primary estimating equation. Column 1 presents the estimates for our base specification, which only controls for county and state-by-year fixed effects. From Panel A, the reduced form estimates imply an effect on population of about 5%, or about a 5.25% effect of NFIP on population when scaling by the first stage.

[10] In the naïve approach, we use actual NFIP take-up instead of map publication to determine treatment status and timing:

$$migration_{cst} = \sum_{l=-\underline{L}}^{\bar{L}-1} \tilde{\delta}_l \cdot newNFIP_{cst-l} + \tilde{\delta}_{\bar{L}} \cdot postNFIP_{cst-\bar{L}} + x_{cst}\tilde{\beta} + \tilde{\lambda}_{st} + \tilde{\gamma}_{cs} + \tilde{\varepsilon}_{cst}$$

In Column 2, we add county, time-varying controls, including per-capita income and unemployment rates. Including these controls does not significantly change our estimate. In Column 3, we add a one year lead of our treatment variable as a falsification check. This directly tests whether counties diverge in population outcomes in the year before NFIP. Consistent with the evidence we showed in Figure A.15, we see no effect prior to treatment for all three outcomes.

We might be concerned that, rather than responding to NFIP availability, households are instead responding directly to previous major disasters which triggered a response from these communities in the form of entry into NFIP. For example, Gallagher (2014) finds an increase in insurance take-up following a flood. In Column 4, we test whether this is a potential confounder by controlling for all water-related nationally declared disasters. Since we obtain similar estimates, we conclude that these disasters do not bias our estimates.

Finally, in Column 5 we introduce county-specific linear time trends. This specification serves as an additional robustness check and will account for any linear trends in unobservables which might be correlated with our instrument and the migration outcomes. Although using this specification causes our point estimates drop for all outcomes, the main estimates are still statistically significant. In addition, they are not statistically different from the estimates we obtain using our primary specification.

Next, we decompose this effect on total population into two sources of variation—residents deciding not to move from one year to the next (non-migrants) and individuals moving into a new county (migrants). These results can be seen in Panels B and C, respectively. Our estimates suggest that most of the effect on population is coming from residents deciding not to move, where the counterfactual—absent NFIP—would have been to move out. After scaling, this amounts to a 5.6% effect of NFIP on non-migrants. Although we estimate a 3% effect of NFIP on in-migration, these coefficients are estimated with less precision than our results for population and non-migrants.

To further verify the robustness of our main results, we perform a series of placebo treatments. For each permutation, we randomly *re-assign* treatment groups (*FHBM*) to each county.[11] Next, we randomly assign each county a time path for *FIRM* assignments.[12] In each permutation, we construct our placebo post-treat variable and estimate Equation 3.2 on each outcome. This procedure provides insight as to whether our estimates are significant just by chance. We perform 1,000 permutations using the same specification as Column 1 of Table A.9—with only county and state-by-year fixed effects. The resulting placebo kernel densities of the placebo estimates are presented in Figure A.16. In each of the panels, the vertical dashed lines correspond to the respective Column 1 estimates in Table A.9. We use this distribution of placebo estimates to compute one-tailed p-values. Panels A and B show that our point estimates are in the right tails of the distribution. In fact, from 1,000 permutations, no placebo estimate is greater than our true point estimates. The results for our in-migration outcome produce a one-sided p-value of 0.015. These results are consistent with our county-level clustered standard errors, and confirm that it is highly unlikely that the size of our estimates are the result of chance.

3.6 Effects of NFIP, by Risk Severity

In this section, we exploit further variation in historical flood risk within treated groups in order to *directly* measure the heterogeneous effects of NFIP on high-risk areas. That is, although FEMA assignment of FHBM is largely driven by the level of flood risk a given community has, there is still significant variation within this level of risk which is not fully accounted for in our main approach. Therefore, we implement a triple-diff approach in order to estimate the additional impact of NFIP on areas which experienced *historically* more floods than other areas that are also treated.

[11] To avoid restrictions on the distribution of FHBM between zero and one, we simply re-assign values for each county directly from the sample of counties.

[12] That is, because each community in a county is treated at different times, we observe unique paths of *post-FIRM* for each county. Therefore, we randomly take an empirical path for this variable from the sample of counties, and re-assign it as a placebo path.

In Figure A.17, we split our data into below and above median[13] annual flood episodes and present estimates for Equation 3.4 for each subsample. It is clear that the majority of the effect of NFIP comes from the high-risk treated counties in our sample. Similar to our difference-in-differences estimates, we see very little to no divergence between groups for all outcomes. In addition, the divergence in the post-period between high- and low-flood risk areas does not seem to be immediate, in contrast to the average treatment effects in Figure A.15.

The triple-diff estimates from Equation 3.3 are presented in Table A.10. Column 1 presents our base specification, fully controlling for all pair-wise interactions. We include demographic controls in Column 2, and add declared disaster controls in Column 3. The results in Panel A suggest that areas with one additional flood per year have an additional effect of about 3% on population. With an annual flood risk standard deviation of 1.45, these estimates suggest that NFIP has an additional impact of about 4.4 percent for a one standard deviation increase in flood risk. This result fully characterizes the types of incentives which NFIP produces, illustrating that NFIP yields its largest effects in the riskiest of counties.

As with our main results, we decompose the population into the number of non-migrants, in Panel B, and in-migrants, in Panel C. Similar to our results in Table A.9, most of the effect seems to come from the increased propensity of existing residents to stay in risky counties, as opposed to the outside influence of in-migrants. For our primary specification, we estimate significant effects of 3-3.7% on the number of non-migrants. Though we lack statistical power, our estimates still imply a meaningful 1.2-1.6% effect on the number of in-migrants per additional flood.

Assuming flood damages are proportional to population size, our triple-differences estimates suggest that the moral hazard produced by the NFIP has been responsible for significant costs from major historical floods, such as those coming from Hurricanes Katrina and Harvey. Given that

[13]The median annual flood episode is about 0.82.

Orleans Parish (New Orleans), Louisiana ranks in the 75th percentile in historical flood risk (see Figure A.11) in our sample, our estimates suggest that the NFIP contributed to a 6.6% increase in costs attributed to Hurricane Katrina. As for Harris County (Houston), Texas, which ranks outside the 90th percentile in historical flood risk, we estimate that the NFIP was responsible for approximately 15% of the damages from Hurricane Harvey.

Furthermore, these estimates imply that the true level of risk is not being incorporated into efficient insurance rates. As these estimates do not directly disentangle the utility from reduced risk from the disutility from insurance costs, our results imply that premiums are too low in these areas. If policy is intent on providing the right incentives to encourage adaptation to future risks of climate change, it must consider the moral hazard produced by the NFIP.

3.7 Discussion and Conclusion

This paper presents evidence of costly unintended consequences produced by the U.S. National Flood Insurance Program (NFIP). Our findings show that population increases in flood-prone areas as a direct response to community enrollment into the NFIP. This program provides highly subsidized flood insurance, securing households against expensive damages from future floods. Thus, our findings suggest that the benefits households receive in the form of a reduction in potential risks, far exceed their premiums, thereby altering location incentives.

The growth of communities in flood-prone regions of the U.S. produces significant costs following major disasters. This type of behavior has large implications in the midst of climate change and rising sea levels. Shorelines in the U.S. account for only 10% of land area, yet the populations residing there make up nearly 39% of the total U.S. population (National Oceanic and Atomospheric Administration—NOAA). As climate change risks inevitably increase the occurrences of future floods, the population will need to adapt in an effort to mitigate these risks. This may mean developing in less risky areas.

Subsidizing flood losses provides consistent incentives to rebuild and reside in areas with high risk. In the case of national flood insurance, our results indicate that the true level of risk is not being properly incorporated into efficient insurance rates. These inefficient rates produce costly externalities, for example, in the form of increased damages from natural disasters. Our estimates of the moral hazard produced by the NFIP suggest that it may have contributed to nearly 6.6 percent of the damages from Hurricane Katrina, and up to 15 percent of the damages from Hurricane Harvey.

Adaptation may be a necessary component of climate change, as the number of major disasters and flood losses are anticipated to increase (Michel-Kerjan and Kunreuther, 2011). This means accounting for some of the perverse incentives created by subsidized flood insurance. If policy is intent on providing the right incentives to encourage adaptation to future risks of climate change, it must consider the moral hazard produced by the NFIP. With growing concerns of the financial sustainability of the National Flood Insurance Program, this may mean restructuring the program sooner rather than later.

4. DOES THE U.S. PIPELINE INFRASTRUCTURE PROMOTE CLEAN ENERGY USE? EVIDENCE FROM THE EPA'S MERCURY AND AIR TOXICS STANDARDS

4.1 Introduction

The aging coal fleet in America is made up of over 600 plants, generating nearly 2 billion kilowatthours of the nation's electricity annually. Approximately 88 percent of coal-fired capacity was built between 1950 and 1990, and the capacity-weighted average age of operating coal facilities is 39 years (EIA).[1] This preexising fixed capital in place for coal-generated electricity has arguably made large contributions to inertia in fuel source decisions by increasing the opportunity costs of using other fuels. However, recent increases in regulation on coal plants have forced utilities to consider alternatives.

In principle, electric power plants have multiple energy sources to choose from when generating electricity. Historically, coal has been the dominant source of energy. The primary reason is due to its abundance and relatively inexpensive cost. More recently however, the regulatory environment has increased costs on coal users. Natural gas prices have declined significantly in the past decade, making it an attractive alternative. Furthermore, the regulatory environment around natural gas as a source of generating power is much less, considering its low emissions levels.

Recent regulatory actions have lead many operators of old, inefficient coal plants to consider a move to natural gas. These coal-to-natural gas conversions seem to be a feasible alternative to investment into costly emissions reductions equipment. NRG Energy has made the switch at a number of their plants, including four conversions in 2016. NRG spokesman David Gaier commented on these conversions in an Energy News article explaining that most of the existing equipment associated with their coal-fired units required no additional modification, and operated the same

[1] https://www.eia.gov/todayinenergy/detail.php?id=30812

as they did on coal with exception of the fuel source.[2] Gaier further noted that the conversions would produce a projected reduction in emissions of 99.9 percent in sulfur dioxide, 97 percent in particulate matter, 34 percent in nitrogen oxides, plus a carbon dioxide reduction. However, these conversions came at a significant cost to NRG, as it was necessary for the company to build new gas lines to their facilities.

Coal-fired power plants looking to comply with newly imposed regulations must make a choice between investment into costly equipment, shutting down the plant entirely, or as NRG has done with a number of their plants, convert to an alternative, less regulated, fuel source. The infrastructure required for natural gas use imposes additional fixed costs on plants looking to switch. In this paper, I examine the role of the U.S. natural gas pipeline infrastructure in coal plant decisions to convert to clean natural gas energy. Under circumstances in which coal plants are sufficiently close to the preexisting infrastructure, connecting to the network and switching to natural gas may be a worthwhile strategy to comply with emissions reduction mandates. But high construction costs, associated with long distances of pipeline, may make plants further away from the network consider other alternatives.

Even decades after the passage of the 1990 Clean Air Act Amendments, very few restrictions were placed on the emissions of air pollutants by power plants. Coal usage made up nearly 45% of electricity generation as recent as 2010, which imposed costly externalities on the environment. Not only is coal a high emitter of greenhouse gasses—emitting twice as much carbon dioxide (CO2) as natural gas through electricity generation—but, it is also a large contributor to many toxic air pollutants.

Though not directly harmful to humans, burning coal is the largest single source of CO2 emissions from human activity. This makes coal activity one of the main drivers of climate change, as CO2

[2]https://energynews.us/2017/02/24/midwest/conversion-to-natural-gas-brings-new-life-to-aging-coal-plants/

has contributed to most of the earth's warming, out of any other greenhouse gas. Given that the electricity sector makes up about one-third of greenhouse gas emissions, one could easily argue that coal power plants play the largest role.

Aside from greenhouse gases, burning coal also produces many harmful pollutants. Sulfur dioxide and nitrogen oxides lead to acid rain and smog, and are precursors to fine particulates ($PM_{2.5}$). The result is toxins that can spread over a wide area, killing many fish and plants, and are particularly harmful to human health (e.g., Chay and Greenstone, 2003; Chay et al., 2009; Neidell, 2004; Currie and Neidell, 2005; Currie et al., 2003; Chen et al., 2013). Coal is also a polluter of mercury—coal-fired plants making up over half of the mercury pollution from human activity (greenpeace.org). Mercury is a harmful pollutant which has severe effects on human brains and nervous systems.

Finally, coal can release harmful particles into the air, even when standing idle, not being burned. For example, Jha and Muller (2017) find that a 10% increase in coal stockpiles results in a 0.06% increase in the average concentration of $PM_{2.5}$ within a 25 mile radius of a plant. Furthermore, they find that a 10% increase in $PM_{2.5}$ leads to a 1.1% increase in adult mortality, and a 6.6% increase in infant mortality.

In 2011, the U.S. Environmental Protection Agency (EPA) announced standards to limit some of these major toxins being produced by power plants. The Mercury and Air Toxics Standards (MATS) created a regulatory environment with the objective of reducing air pollution from coal- and oil-fired power plants. The final rule set standards for all hazardous air pollutants (HAPs) emitted by coal- and oil-fired electric generating units (EGUs) with a capacity of 25 megawatts or greater (EPA). The standards set technology-based emissions limitations, and anticipates reductions of 90 percent of mercury emissions, 88 percent of acid gases, and 41 percent of sulfur dioxide emissions (EPA).

For the first time, in 2016, natural gas usage by power plants superceded coal usage. This has had large scale benefits in terms of impact on the environment. Though methane in general is a much stronger greenhouse gas than CO2, burning natural gas produces about 50% less CO2 than burning coal (Wang et al., 2011). This makes it a significantly more efficient energy source in terms of greenhouse gas emissions. For these reasons, natural gas is often framed as a potential bridge to a low carbon future.

This surge in natural gas has most likely been driven by multiple factors. The recent hydraulic facturing (fracking) boom has played a large part in the availability of natural gas in the U.S. This surge in supply has reduced prices of natural gas, and has made it more competitive as a fuel source for electricity generation, in terms of cost, with coal (Joskow, 2013; Fell and Kaffine, 2018). The fracking boom has additionally produced meaningful economic benefits (e.g., Muehlenbachs et al., 2015; Hausman and Kellogg, 2015; Feyrer et al., 2017; Bartik et al., 2017).

The EPA has additionally played a major role in this increase in natural gas usage. MATS, along with other regulations, like EPA's proposed Rule 111(d) to limit greenhouse gas emissions, place significant pressure on preexising coal- and oil-fired generating units. Many of the plants under these regulations have existed for decades, and still operate old, inefficient generators. In order to comply with the standards, operators must consider whether adding emissions control equipment is an economical way to bring their inefficient coal- and oil-fired plants into compliance. Additionally, operators may ultimately decide to shut older plants down and bring newer, more efficient plants into operation.

Aside from investing into costly equipment to bring these plants into compliance, operators at many of these older coal plants have recently begun considering fuel conversions. Depending on the plant, converting a coal boiler to fire natural gas may be significantly less expensive than other alternatives. Most of the existing generation equipment for these coal units require no modification, and operate with natural gas just as they would with coal. But the key is whether these plants have the available resources to make these conversions. One main necessity of any coal-to-natural gas conversion is having a constant supply of fuel. Transportation of natural gas by ground must be done in its liquid state, which is considered very cost-inefficient, thus, it is hardly ever used as a means of shipment. Therefore, for a plant to have access to natural gas, they must have access to the pipeline network.

Attaining access to the pipeline network can be a costly venture (BPA, 2004; Oliver, 2015; Smith, 2016). This requires building a lateral transmission line to connect directly from the plant to the mainline network. Not surprising, pipeline mileage is an important and salient component of the costs incurred by a plant. Thus, plants marginally closer to the preexisting network have an advantage. These distances are also an important feature during the permitting process for a pipeline project. The Federal Energy Regulatory Commission (FERC) performs environmental impact analyses, considering the extent of potential damages produced by construction, before granting approval. Not surprising, there is a strong relationship between these distances and permit approval. Plant's presumably take this information into account, adding to the marginal cost of pipeline, and thus, affecting conversion decisions.

Though increased usage of natural gas as a result of the fracking boom and reduced natural gas prices have been well-documented (e.g., Joskow, 2013; Knittel et al., 2015; Brehm, 2017; Fell and Kaffine, 2018), to my knowledge, little has been studied on induced demand from the pipeline infrastructure. In this paper, I examine the effect that the pipeline infrastructure has on a power plant's decision to convert an old coal- or oil-fired boiler to natural gas. I study this decision in the context of a plant subject to MATS compliance, applied at the Electric Generating Unit (EGU)-level. Anecdotally, this is a policy that has had the largest impact on recent coal-to-natural gas conversions.[3] Understanding these additional incentives produced by the pipeline network is important as an alternative policy lever in which policymakers can exploit, or at least consider when granting approval for major pipeline projects. Though there are significant short-run costs associated with pipeline projects (e.g., Clay et al., 2017), such as ecological degradation, emissions from machinery, methane leaks, and in the case of oil pipelines, oil spills, little attention has been paid to the long-run environmental benefits produced. Manipulation of the pipeline infrastructure as a means of inducing old, inefficient coal plants to convert to natural gas may be a viable mechanism in ending the natural tendency toward coal as a primary energy source.

To estimate the impact of pipeline proximity on the plant's propensity to convert a coal- or oil-fired boiler to natural gas when faced with MATS compliance, I implement a triple-differences estimator. Specifically, I examine first the treatment effect of MATS on the treated group—EGUs with capacity of 25 megawatts or greater.[4] This treatment of MATS produces a natural inclination for a power plant to convert to natural gas, as it becomes much more expensive for the plant to operate on coal or oil. I then estimate the heterogeneous effects of MATS based on the plant's pipeline proximity.

I estimate the effect of pipelines at the EGU-level, as this is the level which is treated by the MATS program. As I am primarily interested in a plant-level decision, in order to attain the average effects of a representative plant I employ weighted least squares estimation (WLS) and weight by the inverse number of plants (Solon et al., 2015). This approach allows one to interpret the estimates as an average effect on the plant (rather than EGU), and yields results similiar to my ordinarly least squares (OLS) estimates. Reduced form results demonstrate MATS' significant role in inducing conversions, and that the pipeline infrastructure adds substantially to this effect. As I estimate no effect on oil-fired EGUs, I focus primarily on coal MATS-treated EGUs. Difference-in-differences

[3] https://energynews.us/midwest/conversion-to-natural-gas-brings-new-life-to-aging-coal-plants/

[4] I estimate this relationship within a difference-in-differences (and triple-differences) context, rather than in a regression discontinuity design due to insufficient bandwidth on either side of the threshold, as 25 MW is on the lower end of the distribution of EGU capacities.

estimates suggest that MATS increases the probabilty of an EGU being converted by up to 10 percentage points in 2016, two years following MATS. Triple-difference results suggest that a 10 mile decrease in distance to pipeline network produces a 4 percentage point increase in probability of conversion under MATS. Equivalently, a 10% reduction in pipeline distance produces an additional 6.4% increase in likelihood of conversion for an EGU regulated under MATS.

As my reduced-form estimates only explain the effect of the pipeline infrastructure under the specific lens of the MATS policy, I implement a structural approach which allows me to isolate the incentives produced by ease of access to the network, even in absence of MATS. My dynamic model of the plant's decision to convert from coal to natural gas derives closely from the procedure laid out by Rust (1987), though my setting requires significant modifications. I model the decision at the EGU-level, where the plant observes the two states—MATS regulation (dynamic) and proximity from preexisting pipelines (static)—in every period, and strategically decides whether to convert or keep as coal powered. The decision model allows for co-dependence between generating units, as I assume a lateral pipeline to the plant only has to be built once—when at least one unit is converted. Comparable to the framework of Hotz and Miller (1993), I model the decision to convert as a terminal node, where once the plant has converted, it does not convert back. I consider this a reasonable assumption as no reverse conversions are observed in my data.

This approach fully exploits the MATS program as a means of identifying the marginal effects of pipeline proximity on the plant's fixed cost of conversion. However, inference is not limited to the MATS program as the reduced form estimates are. Similar to the reduced form estimation, fixed costs associated with pipeline proximity are identified exploiting the significant costs imposed on MATS-treated units. When MATS is enforced on a given unit, conversions become a viable alternative to other methods of compliance, such as costly scrubber systems, and thus, proximity to pipelines becomes salient to the coal plant. Imposing that pipeline proximity only affects the plant's decision through linear fixed costs, at the plant-level, allows me to identify these cost parameters separately from MATS parameters, thus, isolating the impact of infrastructure. Cross-checking my model against the reduced form estimates yields near identical results.

For better insight into the potential of infrastructure policy in promoting clean energy, I use my estimated model to perform two counterfactual policies, each simulating pipeline subsidies. First, I examine the counterfactual outcomes of plants subject to MATS under fully subsidized pipeline connection. These subsidies lessen the costs on plants which MATS imposes, and provide overall stronger incentives to adopt cleaner natural gas. Simulation results show a 23% improvement in terms of reduced environmental externalities and an 8% reduction in costs imposed on the coal plant. Furthermore, isolating the effect of pipelines, in absence of MATS, I find that pipeline subsidies alone yield one-third of the conversions that mandates to reduce emissions through MATS do, and produce a present value of $10.1 billion in benefits from reductions in carbon dioxide, sulfur dioxides, and nitrogen oxides. Estimates of reduced damages fully account for secondary sources of pollution produced from these emissions, such as particulate matter. In addition, these results translate into a marginal reduction in damages of $2.6 million per-mile of pipeline.

The remainder of the paper proceeds as follows: In Section 4.2, I discuss the background of the U.S. natural gas pipeline infrastructure, as well as describe the data. In Section 4.3, I discuss my reduced form methodology, and present causal estimates. In Section 4.4, I present the dynamic decision model to be estimated. In Section 4.5, I present model estimation and identification details. In Section 4.6, I evaluate the impact of the pipeline infrastructure by simulating a series of counterfactual infrastructure policies. In Section 4.7, I conclude with a brief discussion of my findings.

4.2 Pipeline Infrastructure & Data

4.2.1 The U.S. Natural Gas Pipeline Infrastructure

The U.S. natural gas pipeline infrastructure is a massive network of 5,256 pipelines, transporting about 6.3 billion cubic feet of natural gas per day (bcfd). This network is the only practical way to transport natural gas to its consumers. The alternative, but costly, method is to ship natural gas in its liquified state. Thus, the pipeline network plays a crucial role in making this resource available for users.

As a potential substitute for coal or oil in electricity generation, natural gas is significantly cleaner. Not only is it cleaner on toxic emissions, but natural gas pollutes approximately 50% less carbon dioxide when burned than that of coal. This makes it a clean alternative source of energy. Though natural gas burns cleaner, when the methane is leaked into the atmosphere, it produces an even stronger greenhouse gas than CO_2. It is not surprising then, that pipelines come with a high degree of scrutiny. It is not uncommon to observe protests against new pipeline projects, in addition to an abundance of media attention. Much of this controversy is warranted, not only because of the potential methane leaks, but the additional greenhouse gases produced from machinery, as well as potential environmental degradation incurred during construction.

There is significant concern that major pipeline projects create additional disturbance in regional ecosystems. Construction comes with excessive noise pollution, as well as air pollution emitted from heavy machinery. Furthermore, building a new pipeline contributes to intensive alteration of geology and terrain. Much of the ongoing debate about the pipeline infrastructure has focused on these costs. Though these (potentially short-run) environmental costs cannot be understated, this paper primarily focuses on the long-run environmental benefits of pipelines.

The U.S. pipeline network is made up of various inter- and intrastate pipelines. These are analogous to inter- and intrastate freeways. A map of the current U.S. pipeline infrastructure is presented in Figure A.21. Within the network are three main types of pipelines: gathering systems, transmission systems, and distributional systems. The natural gas gathering pipeline system gathers unprocessed gas directly from the production wells and delivers it to the processing plant. The raw material is then transported with large transmission pipelines, moving the gas from facilities to ports, refiners, and cities across the county. Finally, the distribution system is made up of the main distribution line, and the service distribution line. The main distribution line moves gas from city to city. The service distribution line system is made up of smaller lines, connecting mainlines into homes, businesses, and power plants. The pipeline spurs connecting directly to consumers, including power plants, are often referred to as hinshaw (or lateral) pipelines.

Similar to residential homes running gas, a power plant that generates electricity from natural gas must be connected to the infrastructure. In this paper, I examine the effect of infrastructure costs, in the form of construction of a lateral pipeline from a plant to the network, and the role these costs play in a coal plant's decision to convert to natural gas. I interpret any positive distance between the coal plant and the network as being proportional to the infrastructure cost the plant faces in order to convert from coal to natural gas. Figure A.20 presents the distribution of these distances. In many cases, this distance is a small fraction of a mile, implying that the plant is already connected, imposing minimal infrastructure costs to convert. Often these distances are several miles, imposing a high cost of conversion. As I will show in this paper, when faced with significant regulatory constraints, these distances play a major role in a coal plant's decision to convert.

4.2.2 Data

The majority of the data in this paper come directly from the U.S. Energy Information Administration (EIA). Unfortunately, there are significant limitations on the availability of pipeline data. As regulator and permiter of the pipeline infrastructure, FERC collects and makes public all proposed, and approved, pipeline projects. These data include detailed information on service dates, mileage, capacity, and states in which the projects took place in. However, the most granular geographic information available from these data are the state of operation, which is not sufficient for this project as proximity from pipeline is of primary interest. It has become apparent that national security concerns play a major role in the limited availability of pipeline data.[5]

The pipeline data I use for this project come from an EIA geographic information system (GIS) dataset of the natural gas pipeline infrastructure, published in 2012.[6] This dataset contains precise geographic information for all U.S. interstate and intrastate natural gas pipelines. These GIS data can be directly observed from the maps in Figure A.21.

Data on power plants are collected from EIA form 860s. These data are recorded at the electric generating unit (EGU)-level for power plants with 1 megawatt or greater of combined nameplate capacity. A generating unit may run on multiple energy sources. The most common example is a coal-powered unit that uses natural gas as starter. EIA-860 forms require plants to submit starter fuels as the third or greater source of energy. Energy sources used for electricity generation are listed as the primary and secondary (when applicable) sources. Therefore, in this paper I focus only on a unit's primary or secondary energy source.

From the information in EIA-860s, I am able to extract the total number of units for a power plant in each year, and of primary interest, the number of natural gas- or coal-powered units. EIA also collects exact coordinates of the power plants, which will give me my spacial variation with respect to proximity from the pipeline network. The primary outcome of interest in this paper is natural gas conversions, so the analysis is done at the generating unit-level, as a given coal plant may only

[5]Upon discussions with FERC, I was told that pipeline companies are not required to file GIS data for their projects. Curiously, the data I use for this paper come from a GIS dataset of the pipeline network from the EIA. Up to this date, I have not received an answer from FERC as to where these spacial data ultimately come from. At least one company (MapSearch) collects construction time and spacial data directly from pipeline operators, making these data available to customers at a cost infeasible for this project.

[6]This dataset can be found here: https://hub.arcgis.com/items/f1d3e4fecd56429c9a3bd898d8134d2a. A more recent version can be found here: https://www.eia.gov/maps/layer_info-m.php

convert a fraction of their units. I build my panel dataset from the 2005 cross-section of EGUs, all of which are coal-fired; though a portion of my analysis will be done on oil-fired units as well. The total number of coal plants as of 2005 is 481, and the total number of coal-fired units is 1,144 (see Table A.11). My panel dataset includes years 2006 to 2016, and records whether a unit fuel conversion has taken place. I define an EGU as converted to natural gas if it begins with coal or oil (and not natural gas) in 2005 as its primary or secondary energy source, but records natural gas as its primary or secondary energy source in a period following 2005. This data assumption has been verified with an EIA representative as a reasonable assumption for declaring an EGU as converted.

I examine these natural gas conversion under the lens of EPA's MATS policy, which places significant standards on coal- and oil-fired units of 25 megawatt capacity or greater. The distribution of nameplate capacities for the coal-fired units (of primary interest in this paper) in my sample can be seen in Figure A.19. The majority of coal units are significantly larger than the small, unregulated units. In addition, I split the sample of generating units up by their prime move in Table A.11, which illustrates the majority of conversions (18%) come from steam turbines.

I use measured distance from a power plant to the nearest pipeline segment as a proxy for conversion costs associated with access to natural gas. Though presumably measured with error, one would expect that a plant located far away from the network would have to invest significant capital into the construction of a pipeline connecting them to the network (hinshaw pipeline).[7]. I calculate these distances using EIA data of power plant coordinates and my GIS pipelines data. Although I do not observe pipeline construction dates, I exclude any pipeline segments which show a data published date of 2011 or later. These segments make up less than a tenth of a percent of the entire pipelines dataset. However, if included, I may be concerned that the MATS policy investigated in this paper would have a confounding effect on new pipelines. Furthermore, because of above

[7]An interesting article presenting anecdotal evidence of the high capital costs associated with building pipeline spurs (or hinshaw pipelines) to the mainline transmission for plant natural gas conversions can be found here: https://www.power-eng.com/articles/print/volume-119/issue-6/features/coal-to-gas-plant-conversions-in-the-u-s.html

average distances to network, I drop observations in Hawaii and Alaska. The nearest distance to pipeline network are presented in Table A.11. Distances are measured at the plant-level, however, I partition these observations into EGU-level groups. I observe slightly shorter distances for units subject to MATS, about 9 miles, than units not subject to MATS, about 11 miles. The distribution of these distances can be seen in Figure A.20.

In Section 4.6, I translate converted EGUs in terms of emissions levels. For this exercise, I use plant-level emissions data from the EPA's Emissions & Generation Resource Integrated Database (eGRID), which uses calculated emissions factors to estimate air emissions for five pollutants: nitrogen oxides, sulfur dioxide, carbon dioxide, methane, and nitrous oxide. These data are not available for every year examined in this analysis. The years I will be using include 2004, 2005, 2007, 2009, 2010, 2012, 2014 and 2016. The majority of the data in eGRID is collected from EIA forms 860 and 923.

4.3 Reduced Form Evidence of the Role of Infrastructure

To estimate the role that the natural gas pipeline infrastructure plays in power plant use of natural gas, I exploit quasi-experimental variation in MATS. The MATS program specifically aims to reduce emissions of heavy metals, including mercury, arsenic, chronium, and nickel; and acid gases, including hydrochloric acid and hydrofluoric acid (EPA).[8] These emissions are a byproduct of burning coal, though not significantly related to the buring of natural gas. MATS is enforced on power plants by mandate. For case-specific circumstances, the EPA will grant extensions to bring plants into compliance within one additional year from the enforcement date. Other violations are managed on a case-by-case basis.[9]

Announced in 2011, and officially signed December, 2011, MATS granted plants with EGUs of 25 megawatts (MW) or greater four years to comply with the standards. Power plants affected must

[8] See MATS Fact Sheet: https://www.epa.gov/sites/production/files/2015-11/documents/20111221matssummaryfs.pdf

[9] See: https://www.epa.gov/sites/production/files/documents/mats-erp.pdf

limit their toxic emissions, ultimately preventing 90 percent of the mercury in coal burned at power plants from being emitted into the air (EPA). To comply with the standards, many plants invest in emission reductions equipment, such as scrubber systems, which partially remove particulates and certain gases from exhaust streams. Some operators of old coal plants, might find it most cost efficient to close down the plant and open a new natural gas powered plant. Another option, which has become more popular in recent years, is to convert a regulated coal-fired boiler to fire natural gas.

Although most of the existing generation equipment associated with coal units require no modification, and operate with natural gas just as they would with coal (Energy News Network), plants deciding whether to convert to natural gas must account for the fixed costs associated with conversion. Infrastructure must therefore play a large role in this decision. I estimate the role that infrastructure plays in these coal-to-natural gas conversions, exploiting MATS as an as-good-as-random intervention making pipelines salient to certain plants. Specifically, I implement a triple-difference (DDD) estimator isolating the impact of pipeline proximity under the lens of an as-good-as-random treatment to coal plants—MATS. For my difference-in-differences (DD) estimates of the impact of MATS on conversions, my identification assumes that treated units (≥ 25 MW) would track the control units (< 25 MW) in conversions, absent MATS. Similarly, to isolate the impact of pipeline proximity, I assume that near- and far-proximity MATS-treated units would track each other in conversions, absent MATS. Note that this will only highlight the additional impact of pipeline proximity under the lens of MATS. In Section 4.5, I will estimate a model which generalizes this effect of pipelines beyond the scope of MATS.

As MATS was announced in 2011, but affected plants had up to 4 years to comply, I should expect to see a causal effect from the program beginning prior to full implementation. Ex ante, we might expect plants to begin responding to the program as early as 2011. Therefore, I will estimate a fully dynamic DDD, allowing for deviation from baseline anytime following announcement in

2011, and estimate off of the relative differences prior to announcement. Specifically, I estimate the following equation:

$$
\begin{aligned}
converted_{ijt} = & \sum_{s=2011}^{2016} \delta_s \cdot \mathbb{1}(capacity_{ij} \geq 25\ MW) \times \mathbb{1}(t = s) \\
& + \sum_{s=2011}^{2016} \beta_s \cdot \mathbb{1}(capacity_{ij} \geq 25\ MW) \times \mathbb{1}(t = s) \times distance_i \\
& + \alpha \cdot \mathbb{1}(capacity_{ij} \geq 25\ MW) + \eta_t \cdot distance_i + \lambda_i + \gamma_t + \varepsilon_{ijt}
\end{aligned}
\tag{4.1}
$$

where $\mathbb{1}(\cdot)$ is the indicator function and $\mathbb{1}(capacity_{ij} \geq 25\ MW)$ is an indicator for whether plant i's EGU, j, has nameplate capacity of 25 megawatts or greater. Thus, this indicates whether the EGU will be subject to MATS compliance. $distance_i$ indicates the distance, in 100s of miles, of power plant i to nearest pipeline segment. $\alpha \cdot \mathbb{1}(capacity_{ij} \geq 25\ MW)$ controls for baseline treatment group levels, and $\eta_t \cdot distance_i$ controls for time-by-distance-specific confounders. λ_i and γ_t are plant and year fixed effects, respectively, and ε_{ijt} is the unobserved error. The pipeline infrastructure is defined on a cross-section, prior to the MATS policy to mitigate concerns of endogenous pipeline expansions in areas that have a higher propensity for EGU conversions. Similarly, the EGU sample is defined on a 2005 cross-section due to possible manipulation of capacities following MATS. The coefficients of interest in Equation 4.1, β_s for s equal to 2011 through 2016, estimate the additional effect of pipeline distance on the plants' decision to convert an EGU under MATS compliance in each corresponding year.

My primary estimates from Equation 4.1 come from weighted least squares (WLS) estimation. Estimation is conducted at the unit-level, rather than plant-level, primarily because I am interested in a unit-level decision by the plant, where treatment of MATS is determined on unit-level criteria. Unweighted estimates using ordinary least squares (OLS) will, therefore, place additional emphasis on plants who have more coal units in my sample. As my intent is for the causal estimates to be interpreted as the impact of MATS and the pipeline network on the representative plant's decision, I will be weighting on the inverse number of units in a plant, which will serve to treat each plant equally (Solon et al., 2015).

In estimating Equation 4.1, it's important to note that I exclude any units from the control group owned by plants who also own treated units. Including these in the control group would most likely attenuate my estimates, as plants with large *and* small units may decided to convert their small units, conditional on the large units being converted due to decreasing marginal costs. Thus, these control units would be affected by the intervention as well. I do not exclude these units from the structural estimation in Section 4.5, as my model will account for this decreasing marginal cost of conversion.

The fraction of converted EGUs for the treated and control are plotted over time in Figure A.22. Note that original fuel types are defined in 2005, thus, the figure begins plotting conversions beginning in 2006. As MATS applies to both coal and oil units, I plot conversions for both fuel types. It's immediately apparent that oil units have no obvious effect, thus, I will strictly be examining the impact on coal units for the remainder of the paper. This figure provides strong support of the parallel trajectories assumption of a DD, as in the periods prior to 2011 the treated and control units track each other very closely in conversions.

Figure A.23 plots the treated coal units separately for near- and far-proximity plants. As expected, plants near to the pipeline network begin converting faster, and at a higher rate than plants further away from the pipeline. Note that the response from the plants further from the pipeline network have a slower response, possibly due to permitting and construction times. Unfortunately, I will not be able to account for this delayed response, as I do not observe when the actual decision and investment was made. As the majority of these projects take under two years, I do not expect this to alter my estimates significantly, but only produce a partially lagged effect.

Table A.12 presents the OLS estimates of Equation 4.1, and Table A.13 presents the WLS estimates. Columns 1-2 present the DD estimates of the treatment effect of MATS on the probability of EGU conversion from coal to natural gas for years 2011 through 2016.[10] Estimates suggest that MATS increases the probabilty of an EGU being converted by up to 10 percentage points in 2016.

Columns 3 through 5 list the estimates for my DDD estimates. These estimates suggest that pipeline infrastructure has a significant impact on the power plants' decision to convert when under MATS compliance. These estimates do not change significantly after controlling for year-specific effects of distance in Column 5.[11] Given the WLS estimates in Table A.13, these results suggest that a 10 mile decrease in distance to pipeline network produces a 4% increase in probability of conversion under MATS. With a baseline average distance to pipeline of 10.29 miles and baseline conversion of 0.0648, these results produce an elasticity of -0.635. That is, a 10% reduction in pipeline distance produces an additional 6.4% increase in likelihood of conversion for an EGU regulated under MATS.

4.4 Model

In this section, I present a model of the plant's strategic decision to convert a coal-fired EGU to fire natural gas. By imposing minimal structure on my data, I will be able to uncover the representative plant's marginal propensity to convert from a change in pipeline proximity, even in absence of MATS. That is, whereas my estimates in Section 4.3 only highlighted the additional impact of the pipeline infrastructure on conversions for units subject to MATS, by implementing a DDD approach, estimates from my model will allow me to predict how the network not only complements,

[10]Estimates do not change for the first 4 decimal places when including plant fixed effects, going from column 1 through 2. Note that this is not an error. Any fixed non-time-varying confounders are sufficiently controlled for in a *treatment group* indicator.

[11]The fact that the reduced form estimates are not sensitive to inclusion of fixed effects and additional interactions is important for structural estimation performed in Section 4.5, as the inclusion of these is not computationally feasible.

but may substitute for MATS.

Suppose a representative coal power plant, i, from set of plants N, has profits in period t according to:

$$
\begin{aligned}
\Pi_{it} =& p_t \cdot \sum_{j \in J_i} q_j(f_{jt}) - \sum_{j \in J_i} \mathrm{c}_j(f_{jt}) \\
& - \kappa_i \cdot \mathbb{1}\big(f_{jt-1} = coal \; \forall j \in J_i \;\; \wedge \;\; f_{jt} = ng, \textit{ for some } \; j \in J_i\big) + \epsilon_{it}
\end{aligned}
\tag{4.2}
$$

where $q_j(f_{jt})$ defines the quantity of electricity produced for a given EGU, j, in plant i's set of EGUs, J_i, given the fuel type, f_{jt}, of that EGU. In this model, fuel type is either coal or natural gas (i.e., $f_{jt} \in \{coal, ng\}$). p defines the price paid by consumers of electricity, and thus, the first term is the plant's (conditional) expected revenues from total electricity generation in period t. Similarly, the second term defines the plant's total cost of electricity generation as a function, $\mathrm{c}(\cdot)$, of each generator's fuel type. κ_i defines the plant-specific fixed cost of converting to natural gas. I assume that this fixed cost is only incurred when converting from all coal EGUs ($f_{jt-1} = coal \; \forall j \in J_i$) to at least one natural gas EGU ($f_{jt} = ng$, *for some* $j \in J_i$). I will interpret this cost as an infrastructure cost incurred when connecting to the natural gas pipeline network. Finally, ϵ_{it} is an idosyncratic shock to the plant's profits.

In this paper, I model the plant's decision to convert at the generator level. Therefore, the connection cost, κ_i, is only incurred if the plant is not already connected (defined as an EGU already converted to natural gas). Given the per-period cost structure in Equation 4.2, the marginal cost of the plant converting a single EGU, j, is defined by *Proposition 1*.

Proposition 1: for plant $i \in N$, fixed cost of connection, k_i, and set of fuel types, $\{f_{jt}\}_{j \in J_i}$, the marginal cost of converting a unit $j \in J_i$ at time t is defined as:

$$\begin{aligned} MC_{ijt} = & - \big(\mathrm{c}_j(ng) - \mathrm{c}_j(coal)\big) \\ & \quad - \kappa_i \cdot \mathbb{1}\big(f_{j't-1} = coal \; \forall j' \in J_i \quad \wedge \quad f_{j't} = coal \; \forall j' \neq j, \; j' \in J_i\big) + \epsilon_{ijt}^{ng} - \epsilon_{ijt}^{coal} \end{aligned} \tag{4.3}$$

proof: See Appendix B.10.

Proposition 1 demonstrates that fixed costs only enter the marginal cost of converting j if no other EGU is converted ($f_{j't} = coal \; \forall j' \neq j, \; j' \in J_i$). Conditional on the plant's set of preexisting fuel types, $\mathcal{F}_{it-1} \equiv \{f_{jt-1}\}_{j \in J_i}$ and fixed cost, κ_i, from *Proposition 1* we can derive the expected marginal cost of converting unit j is as:

$$\begin{aligned} \mathbb{E}[MC_{ijt} | \mathcal{F}_{it-1}, \kappa_i] = & - \big(\mathrm{c}_j(ng) - \mathrm{c}_j(coal)\big) \\ & - \kappa_i \cdot \mathbb{1}\big(f_{j't-1} = coal \; \forall j' \in J_i\big) \prod_{j' \in J_i \backslash j} Pr(\neg convert_{ij't} | \mathcal{F}_{it-1}, \kappa_i) \\ = & - \big(\mathrm{c}_j(ng) - \mathrm{c}_j(coal)\big) - \kappa_i \cdot \tilde{\mathbb{P}}_{ijt} \end{aligned} \tag{4.4}$$

The final term defines the plant's expected fixed costs associated with converting j, conditional on the set of prior fuel types and fixed cost of connection. Thus, the term interacting fixed costs is the probability of no other unit being converted, conditional on information at the beginning of t. Notice that, conditional on prior period fuel types, $\mathcal{F}_{it-1}$, whether a prior conversion has taken place is known with certainty (e.g., $\mathbb{1}\big(f_{j't-1} = coal \; \forall j'\big)$), whereas the likelihood of no other conversion taken place is in expectation. One should avoid thinking about this term as the plant's beliefs of converting other units, but rather, the proportion of the fixed cost associated with converting j, conditional on observables. Equation 4.3 is discrete in fixed costs, whereas Equation 4.4 is smooth. Therefore, this term accounts for the decreasing marginal cost of converting a unit, and presents joint dependence amongst units in a plant's fleet.

Given this construct, we can write the plant's unconditional marginal costs as its conditional marginal costs, plus an unobserved (to the researcher) term, μ_{ijt}. That is, plant i's marginal cost of converting unit j is:

$$MC_{ijt} = \mathbb{E}[MC_{ijt}|\mathcal{F}_{it-1}, \kappa_i] + \mu_{ijt} \tag{4.5}$$

According to Equation 4.2, we can write period t's marginal revenue of converting j from coal to natural gas as:

$$\begin{aligned} MR_{ijt} =& p \cdot (q_j(ng) - q_j(coal)) + r^{ng}_{ijt} - r^{coal}_{ijt} \\ =& p_t \cdot (q_j(ng) - q_j(coal)) + r_{ijt} \end{aligned} \tag{4.6}$$

where r^{ng}_{ijt} and r^{coal}_{ijt} denote unobserved revenues associated with unit j using natural gas and coal, respectively, and r_{ijt} combines these terms. I assume that plant i converts unit j if its expected present value of marginal revenues are greater than its expected present value marginal costs of converting. That is, i converts j when:

$$\mathbb{E}_{it}\,\mathbf{PV}\big(MR_{ij}\big) = \sum_{s=0}^{\infty} \beta^s \cdot \mathbb{E}_{it} MR_{ijt+s} > \sum_{s=0}^{\infty} \beta^s \cdot \mathbb{E}_{it} MC_{ijt+s} = \mathbb{E}_{it}\,\mathbf{PV}\big(MC_{ij}\big) \tag{4.7}$$

where $\mathbb{E}_{it}$ is the plant-specific expectation operator for period t, and β is the discount factor, common to all plants. We can decompose the marginal costs and marginal revenues of Equation 4.7 into their coal- and natural gas-specific components and write the plant's Bellman equation for coal generator, j, as:

$$\begin{aligned} V^0_{ijt} = \max \Big\{ & p_t \cdot q_j(coal) - \mathrm{c}_j(coal) + \varepsilon^0_{ijt} + \beta \cdot \mathbb{E}_{it} V^0_{ijt+1}, \\ & p_t \cdot q_j(ng) - \mathrm{c}_j(ng) - \kappa_i \cdot \tilde{\mathbb{P}}_{ijt} + \varepsilon^1_{ijt} + \beta \cdot \mathbb{E}_{it} V^1_{ijt+1} \Big\} \\ V^1_{ijt} = & p_t \cdot q_j(ng) - \mathrm{c}_j(ng) + \varepsilon^1_{ijt} + \mathbb{E}_{it} V^1_{ijt+1} \end{aligned} \tag{4.8}$$

where ε^0 and ε^1 denote the unobserved (to the researcher) components of the marginal costs and marginal revenue associated with coal and natural gas operation, respectively. V^1_{ijt} denotes the Bellman equation for sequential periods following conversion to natural gas, where fixed costs are sunk in future periods.

4.5 Estimation

Fuel costs and electricity prices are not observed in my data for the counterfactual decision of the plant. Furthermore, additional assumptions must be imposed on the plant's expectations, and endogeneity concerns must be addressed if included in the model. My simplification, which does not include prices, should not significantly alter my estimates of interest, mainly, the plant's response to MATS and pipeline proximity. Consider the following simplification of the Bellman equations in Equation 4.8:

$$\begin{aligned} V^0_{ijt} &= \max \left\{ \delta^0_{ijt} + \varepsilon^0_{ijt}, \;\; \delta^1_{ijt} + \varepsilon^1_{ijt} \right\} \\ &= \max \left\{ -c^{coal}_{jt} + \varepsilon^0_{ijt} + \beta \cdot \mathbb{E}_{it} V^0_{ijt+1}, \;\; -c^{ng}_{jt} - \kappa_i \cdot \tilde{\mathbb{P}}_{ijt} + \lambda_t + \varepsilon^1_{ijt} + \beta \cdot \mathbb{E}_{it} V^1_{ijt+1} \right\} \\ V^1_{ijt} &= -c^{ng}_{jt} + \varepsilon^1_{ijt} + \mathbb{E}_{it} V^1_{ijt+1} \end{aligned} \tag{4.9}$$

I model the decision to convert, as a terminal node decision. That is, once the plant has converted to natural gas, they do not convert back to coal. I do not observe reverse conversions in my data.

Thus, in estimating this model, I trim observations following a conversion. In this model, unobserved (to the researcher) revenues are captured in idosyncratic shocks to profits, ε^0 and ε^1. I assume these terms are distributed independently, conditional on observables. In this specification, I allow for period-specific fixed costs shared by all plants, λ_t, included, without loss of generality, on the convert side. The state variables of interest in my model of the plant's decision are the MATS policy and pipeline proximity. The model assumes that MATS affects treated plants' per-period, unit-level costs associated with using coal. It is important to note that MATS may impact fixed costs of the power plants, for instance investment into scrubber systems—used to remove certain particulates or gases from the exhaust of coal-fired plants. In this case, my estimates will represent a period-unit-specific proportion of these fixed costs.

I specify per-period plant costs as a function of the MATS policy, and allow for a high-capacity (> 25 MW) unit specific term. Furthermore, as independent variation in unit-level costs are not observed for natural gas units (especially for the counterfactual decision to convert), I normalize natural gas per-period costs to zero. This normalizes operating costs in the Bellman equation for converting to zero, and thus, the conversion side only depends on fixed costs.

$$\begin{aligned} c_{jt}^{ng} &= 0 \\ c_{jt}^{coal} &= c_{jt} = c_0 \cdot \mathbb{1}(capacity_j \geq 25) + c_1 \cdot \mathbb{1}(capacity_j \geq 25) \times \mathbb{1}(t \geq 2015) \end{aligned} \tag{4.10}$$

where units greater than 25 megawatts are treated with the MATS policy, whereas compliance for smaller units is not required. c_1 is the additional cost imposed on a coal unit for being regulated under MATS, following the compliance period of 2015. This specification is analogous to my difference-in-differences approach, where period-specific fixed costs absorb time-specific confounders.

I model the fixed cost of conversion as a linear function of distance to pipeline. I treat proximity to pipeline as a static attribute for two reasons. First, and most hindering, very limited data are available for the pipeline infrastructure due to national security concerns, as described in Section 4.2.2. Thus, I do not have access to a set of annual pipeline maps to observe the evolution of the network. Second, if access to annual pipeline data were available, I would be concerned that future pipeline projects might be endogenous to plants' propensity to convert, as well as the MATS policy. As the majority of the pipeline network was installed in the 1950s, I don't anticipate estimates to be sensitive to the use of the static infrastructure, and intuitively, a positive distance in my data between a plant and the network which no longer exists should only result in an attenuated estimate of fixed costs.[12] My specification assumes the cost of connecting to the pipeline network is proportional to a plant's distance to the nearest node.

$$\kappa_i = \kappa \cdot distance_i \tag{4.11}$$

Given the parameters of the fully specified model, the non-conversion probabilities, $\tilde{\mathbb{P}}_j$, are endogenously determined for each j; though I must first specify the structure of the plant's expectation of future costs. The static pipeline infrastructure and salient roll out of MATS makes for a simple (and credible) structure for these expectations. As the only time-varying state variable is MATS, I assume that plants expect no change in expected costs in periods prior to the announcement of the policy in 2011. This assumes that plants have myopia with respect to the occurence of MATS, prior to the announcement date. Following the announcement, I assume that plants have perfect foresight with respect to the costs they will incur (if treated) in 2015—the compliance date. For periods following full implementation of MATS (2015), I assume plants have myopia, which, with respect to the states, is equivalent to perfect foresight as long as the policy continues to be implemented. Specifically, I assume the following structure on plant i's expectation of future costs

[12] Though I only cut a tenth of a percent of observations, when conditioning on data published prior to 2011 (announcement date), visually, the 2016 natural gas pipeline network map appears very similar to Figure A.21, suggesting very minimal change has taken place. Most interstate projects that have taken place have expanded interstate capacity, and may not have necessarily changed plant proximities. See EIA pipelines map: https://www.eia.gov/naturalgas/archive/analysis_publications/ngpipeline/index.html

in period t:

Plant Expectations Assumption: for a plant $i \in N$, unit $j \in J_i$, and for some $s > 0$,

$$\mathbb{E}_{it} c_{ijt+s} = \begin{cases} c_{ijt+s} & \text{if } \ t \geq 2011 \\ c_{ijt} & \text{otherwise} \end{cases}$$

As MATS is in full implementation in 2015, the state does not change in following periods. Therefore, it holds that for periods following 2015, $\mathbb{E}_{it} c_{ijt+s} = c_{ijt+s} = c_{ijt}$, for all $s > 0$. Calculating the Bellman equation is now straightforward. Assuming idiosyncratic shocks to profits are distributed independently from the extreme value type I distribution, the value function has the following closed form solution, unconditional on these shocks.

$$V^0_{\sigma,ijt} = log\big\{exp(\delta^0_{ijt}) + exp(\delta^1_{ijt})\big\} \tag{4.12}$$

where δ^0_{ijt} and δ^1_{ijt} are as described in Equation 4.9, $log(\cdot)$ is the natural logarithmic function, and $exp(\cdot)$ is the exponential function. Given the set of parameters, and a first guess for $\tilde{\mathbb{P}}_j$, for periods prior to 2011 and periods following and including 2015, I calculate V^0_σ using value function iteration, under myopia. This is simply a fixed point exercise. For periods equal and greater than 2011, but prior to 2015, I first calculate the Bellman equation in 2015 by value function iteration, then solve the Bellman equation for the period of interest by backwards induction.

Given the initial guess for $\tilde{\mathbb{P}}$, and calculated V^0_σ, I form the conversion probabilities for plant i, unit j, in period t as:

$$\Pr(convert_{ijt}|X_{it},\Theta) = \Pr(\delta^0_{ijt} < \delta^1_{ijt}) = \frac{exp(\delta^1_{ijt})}{exp(\delta^0_{ijt}) + exp(\delta^1_{ijt})} \tag{4.13}$$

where X_{it} are the state variables for all units in plant i and Θ is the set of parameters. Thus, with an initial guess for $\tilde{\mathbb{P}}$, I calculate conversion probabilities, and update the product of non-conversion probabilities according to:

$$\tilde{\mathbb{P}}_{ijt} = \mathbb{1}\big(f_{j't-1} = coal\ \forall j' \in J_i\big) \prod_{j' \in J_i \setminus j} \Pr(\delta^0_{ijt} > \delta^1_{ijt}) \tag{4.14}$$

I then repeat this process until the the values from Equation 4.14 converge. Note that this process is trivial in periods following a single conversion. $\tilde{\mathbb{P}}_{ijt}$ simply becomes zero (because $\mathbb{1}\big(f_{j't-1} = coal\ \forall j' \in J_i\big) = 0$), placing no additional cost of connection costs on the plant. Once the process has converged, I form the following likelihood values:

$$\Pr(convert_{ijt})^{y_{ijt}}\big(1 - \Pr(convert_{ijt})\big)^{1-y_{ijt}} \tag{4.15}$$

where y_{ijt} equals one if a conversion is observed in my data and zero otherwise. My final estimated parameters maximize the joint likelihood across all observations. I will assume that all plants discount future periods at the same rate, with a discount factor of 0.95. Similar to my reduced form estimates, I estimate the model unweighted, as well as weighted by the inverse of the number of units in a plant. This treats every plant equally, rather than placing additional emphasis on plants owning many units, and better approximates the actions of a representative plant.[13]

[13] My weighted estimates weight each likelihood value by the inverse number of generators in the plant's fleet, G_i. Specifically, I maximize the following log-likelihood function: $\sum_{ijt} \frac{1}{G_i} log(\ell_{ijt})$, where $\sum_{ijt}$ represents a sum across plant, i, unit, j, and year, t, and ℓ_{ijt} is the likelihood value from Equation 4.15. Note that this is equivalent to maximizing the product of the geometric means of likelihoods, at the plant-level.

4.5.1 Identification

The final model to be estimated, under the preceding framework laid out, simplifies to the following:

$$\begin{aligned} V_{ijt}^0 &= \max \left\{ \delta_{ijt}^0 + \varepsilon_{ijt}^0, \quad \delta_{ijt}^1 + \varepsilon_{ijt}^1 \right\} \\ &= \max \Bigg\{ -c_0 \cdot treat_i - c_1 \cdot treat_i \times post2015_t + \varepsilon_{ijt}^0 + \beta \cdot \mathbb{E}_{it} V_{ijt+1}^0, \\ &\qquad -\kappa \cdot distance_i \cdot \tilde{\mathbb{P}}_{ijt} + \lambda_t + \varepsilon_{ijt}^1 \Bigg\} \end{aligned} \tag{4.16}$$

Identification of the parameters in this model has similar intuition to the reduced form estimation in Section 4.3. For identification of the costs associated with the MATS policies, I assume that, given the structure of the model, costs of running coal would have followed similar trajectories to the low capacity ($<$ 25 MW) units absent the MATS intervention. Similar to my DD strategy, I control for time-specific costs, λ_t—which map into differences in conversions over time—as well as treatment group-specific costs, c_0—allowing high capacity units ($\geq$ 25 MW) to have different baseline costs, and thus, different conversions rates.

Identification of plant fixed cost of connecting to the pipeline network also comes from similar variation to that described in Section 4.3. In my DDD estimates, I exploit the MATS program as an as-good-as-random intervention which makes pipeline proximity salient. In my reduced form estimates, responses to this proximity to network are only identified for treated units, following implementation (or announcement) of MATS. Baseline responses are not identified in my DDD approach, but rather, are absorbed into group-by-distance, and time-by-distance controls. This model, though requiring additional structure, identifies this baseline response to proximity.

Intuitively, similar to my DDD, pipeline proximity is made salient under MATS. That is, identification of κ comes from the additional shock to coal costs, which makes δ^0_{ijt} in Equation 4.16 large under MATS. Converting becomes an alternative option for compliance, and the plant now considers its proximity. Imposing the assumption that proximity to pipeline enters plant payoffs linear in a *plant-level* (through an interaction with $\tilde{\mathbb{P}}$) fixed cost, identifies these baseline responses to infrastructure, allowing us to interpret these effects beyond the scope of MATS. This construct assumes that the interaction terms between distance and MATS, in Equation 4.1, derive directly from the cross derivatives of MATS and proximity, conditional on MATS occuring, in the conversion probabilities evaluated at sequential years. Thus, the DDD estimates derive from the non-linearity of the model. In Section 4.6, I use my estimated model to gauge how closely it predicts the reduced form estimates in Section 4.3.

4.5.2 Model Estimates

The logit maximum likelihood (MLE) estimates are presented in Table A.14. The costs parameters in the table are presented in positives, where the cost functions they produce are subtracted in the value function, as in Equation 4.9. The coefficient on the treatment group indicator ($\mathbb{1}(capacity \geq 25)$) should not be interpreted causally, but rather, as a term which absorbs confounders in the cost equation correlated with large units (for both coal and natural gas units). Thus, this term may not only absorb costs of operation, but also non-identified constant fixed costs of conversions, specific to unit size. The negative sign on the treatment group indicator, which is subtracted in the value function, suggests that, on average, the MATS-treated group has a lower rate of conversions. This may be due to the latter described higher fixed cost of conversion.

The coefficient on the interaction term, identifying the per-period change in costs of operating a coal-fired generating unit under MATS, suggests a higher cost of operation when under MATS regulations. Identification of this increased *operating* cost assumes that MATS does not increase *conversion* costs for treated units. Finally, the estimated marginal effect of proximity to pipeline is defined by κ. As expected, the sign of this coefficient is positive (which is subtracted in the value function), indicating a higher cost of conversion for plants located far away from the infrastructure.

These estimates are presented for the standard maximum likelihood, as well as weighted maximum likelihood (my primary specification), which weights each observation by one over the number of units in the plant. Each specification includes year-specific fixed costs which absorb any time-specific confounders to costs.

4.6 Evaluating Counterfactual Infrastructure Policies

In this section, I examine the implications of the pipeline infrastructure more thoroughly using the estimated parameters of the model. I simulate the effect of various alternative policies by manipulating the inputs of the model, holding constant other factors. Specifically, I manipulate the MATS policy—where I assert that the coefficient for MATS costs is identified—as well as the pipeline infrastructure—where I assert that the coefficient on proximity is identified. Other factors, whose parameters may not be identified, will be held constant.

I will examine two counterfactual policies, in addition to an ex-post analysis of the MATS policy. In each counterfactual policy, I manipulate pipeline costs to examine the relevance of infrastructure in energy choices. Connection to the pipeline infrastructure imposes intensive costs on the plant in terms of permitting, construction, physical constraints on routes, and other regulatory obstacles. To the extent to which these costs vary with proximity to pipeline, I can examine conversions under varying marginal costs per mile of pipeline. Thus, I examine the extreme case of fully subsidized pipelines in order to isolate the potential benefits which would be produced under zero marginal constraints on pipeline mileage.

My first counterfactual policy estimates the potential benefits of MATS that would have occurred under a subsidized pipeline infrastructure, that produces zero marginal cost of pipeline connection to the plant. Therefore, this combines two policy levers to generate larger benefits than produced under MATS alone. My second counterfactual policy exploits the marginal effect of pipeline proximity identified by my model. Whereas the marginal effect of pipeline proximity on conversions is only identified under MATS in my reduced form estimates, my structural estimation procedure allows me to disentangle these two forces, by asserting that pipeline proximity only enters through *plant-level* fixed costs. Thus, I estimate the impact of fully subsidized pipeline infrastructure under the scenario of no MATS.

My model is specified to produce the optimal decisions by the plant to convert from coal to natural gas. Thus, the environmental benefits from this action are not a direct outcome of the model. In order to provide an interpretation of the environmental benefits of these actions, I make use of plant-level emissions data provided by EPA's eGRID system. Specifically, I will estimate the following regression to provide insight into my model predictions:

$$\textit{Emissions-per-MW Capacity}_{it} = \delta \textit{converted fraction}_{it} + X_{it}\beta + \gamma_i + \lambda_t + \epsilon_{it} \tag{4.17}$$

where *Emissions-per-MW Capacity*$_{it}$ is a measure of plant i's emissions per megawatt unit of capacity, δ estimates the marginal effect of conversion on emissions. As my emissions data are at the plant-level, I aggregate unit-level conversions into fractions of a plant's capacity (holding capacity constant at 2005 levels)[14] converted from coal to natural gas. In the data, I observe 79% conversions in the first period in which I observe at least one conversion. I control for plant-specific factors, γ_i, and year-specific factors, λ_t. I test for potential (observable) confounders by including a set of covariates, X_{it}. For example, one might suspect that plant conversions are correlated with a reduction in annual generation. In this case, controlling for generation might alter my estimates.

Table A.15 presents the OLS estimates of Equation 4.17. Standard errors are clustered on plant. I

[14]My estimates are not sensitive to whether I use baseline 2005 capacity versus time-varying plant capacity, as capacity does not vary significantly over time in my data.

estimate the effect of conversions on three different sources of emissions. Panel A examines the effect of plant conversions on greenhouse gases, using carbon dioxide (CO_2) equivalents. This metric includes plant carbon dioxide, methane (CH_4), and nitrous oxides (N_2O), weighted by the global warming potential (GWP) for the Intergovernmental Panel on Climate Change (IPCC) fifth (2014) (AR5) assessments. In addition, I examine the effect of conversions on two sources of local pollutants reported by eGRID—mono-nitrogen oxides (NO_X) and sulfur dioxide (SO_2). Though eGRID data do not include information on plant direct emissions of particulate matter, these two pollutants are both precursors to $PM_{2.5}$, and therefore, these estimates should contain derivable effects on local particulate pollution.

Column 1 presents the results including only plant and year fixed effects. In Column 2, I add in state-by-year fixed effects, and add additional controls in Column 3. These estimates do not significantly change from Column 1 to Column 3, and the inclusion of plant net generation or plant heart rate (a measure of efficiency) does not significantly effect my estimates. This result argues that plants are not altering their generation (or efficiency) following a conversion in a manner that would also affect emissions levels. This is an important result, as now I can examine the total environmental benefits of these conversions across plants without significant concern of increases in generation from other coal plants countering these benefits. I test this relationship between conversions and generation levels more directly in Appendix B.11.

Greenhouse gases emitted by power plants can have wide spread effects on climate change long term. There are various models that are used to estimate the cost that carbon imposes on the environment. Conventional estimates place this value at about $40 per metric ton of CO_2 in today's dollars.[15] Therefore, I will use this social cost of carbon (SCC) to interpret the impact of my counterfactual policies. In contrast to CO_2, SO_2 and NO_X pollution have a much less trivial effect on the environment, as these pollutants only have localized effects. The external costs on local

[15]See https://theconversation.com/curbing-climate-change-has-a-dollar-value-heres-how-and-why-we-measure-it-70882

pollutants have been well-documented in the literature (e.g., Chay and Greenstone, 2003; Brauner et al., 2007; Anderson, 2016; Clay et al., 2016; Jha and Muller, 2017). To translate these costs into dollar terms, for direct comparisons, I rely on an integrated assessment air pollution model—the Air Pollution Emission Experiments and Policy (APEEP) analysis model (Muller and Mendelsohn, 2007; Muller and Mendelsohn, 2009).

The APEEP calculates the marginal damages of emissions for nearly 10,000 distinct sources of air pollution in the U.S. To do so, a baseline level of pollution is first established for a given county. APEEP then adds one ton of a specific pollutant to that county, and damages are recalculated. The difference between this new estimate and the original baseline damages is the marginal damage for that county. Important for this analysis, this approach captures the damages from secondary pollutants, which are attributed back to the emission that caused them. Therefore, the marginal damage of particulate matter pollution produced from observed SO_2 and NO_X in my data will be contained in these estimates. Muller and Mendelsohn (2007) and Muller and Mendelsohn (2009) translate these damages into dollar costs, aggregating the impacts on mortality, morbidity, agriculture, and various other local damages. I use these estimated county-level marginal damages, along with estimates from Table A.15 (Column 3), to evaluate the external benefits of conversions at the plant-level.

As none of my state variables are denominated in dollars, I can not *directly* make surplus calculations. However, for a basis of comparison, I can pull estimates from the literature and make these calculations indirectly in a very *back-of-the-envelope* manner. In particular, I will translate estimated logit-payoffs into dollars using a conversion of pipeline mileage to dollar costs. That is, given an estimate of cost-per-mile of pipeline (c_p), I can approximate how my logit-payoffs, V, translate into dollars. This is because, given a constant marginal cost of pipeline mileage:

$$c_p = \frac{\frac{\partial V}{\partial distance}}{\frac{\partial V}{\partial \pi}} = \frac{\kappa}{\frac{\partial V}{\partial \pi}} \tag{4.18}$$

where π is the value of (unobserved) profits in dollars, and V is the logit-value of payoffs. There is significant heterogeneity in pipeline costs, so these estimates should be interpreted with caution. I use a ballpark estimate of \$1 million per mile, which anecdotally appears consistent with, though on the lower end of empirical estimates (Oliver, 2015; Smith, 2016).[16] This translation into dollars ultimately amounts to multiplying estimated logit-payoffs by c_p/κ. Thus, reported producer surplus estimates can easily be adjusted by the reader for a different estimate of *price-per-mile*, c_p.[17]

4.6.1 The MATS Policy

In Section 4.3, I analyzed the ex-post effects of the MATS policy and the additional role the pipeline infrastructure plays on coal-to-natural gas conversions in a reduced form setting. I can use those estimates as a benchmark in assessing how well my model predicts in-sample behavior. In Figure A.24, I present the fitted values of the model in a setting analogous to Figure A.23. I split the plants into below and above median proximity from pipeline, simulate the behavior of each unit-plant, and average the model results across plants. The result is very similar to the sample moments in Figure A.23. This result suggests not only that I am identifying similar marginal effects on MATS and pipeline proximity in my structural model, but also that my assumption on expectations are plausibly correct. The initial increase in conversions, beginning in 2011, come from backwards inducting perfectly forecasted MATS-compliance costs to be incurred in 2015. This modeling as-

[16] See https://www.ogj.com/articles/print/volume-114/issue-9/special-report-pipeline-economics/natural-gas-pipeline-profits-construction-both-up.html, Table 6 for a range of estimates of average per-mile cost of pipeline. These figures are reported for different years, as well as different diameters. I consider diameters less than 16 inches, as lateral pipelines to natural gas power plants tend to be between 6 and 16 inches (http://naturalgas.org/naturalgas/transport/).

[17] E.g., if the reader suspects cost-per-mile to be closer to \$2 million, simply multiply my estimates by 2.

sumption seems to proxy well for the true underlying behavior of the plants.[18]

Not accounted for in the model is the permitting and construction time for pipeline projects. As I only observe when a plant converted, and not when construction or permitting began, I am not able to account for this delay. The result of this assumption is an immediate response from plants far away from the pipeline, where in reality, this response is actually lagged. This delayed response should not affect the ultimate long-run outcomes of my simulations, as these differences only depend on the length of time examined. That is, if the data were aggregated into two year increments—well within the length of time necessary for a majority of pipeline projects—a delayed response would not be observed.[19] Thus, for the entire time-horizon of interest, these differences should not be of first order concern.

I present the model estimates in Figure A.24 for both the unweighted and weighted MLE estimates. Though scaled slightly different, the marginal effects appear to be very similar for both estimates. For the remainder of the model simulations, I will use the weighted MLE estimates, as I believe these to be more representative of the average plant.

Table A.16 lists in-sample model prediction of the MATS policy. The column labled Δ*Converted* presents the period-specific effects of MATS on the probability of being converted—the cumulative effects on conversions.[20] These estimates calculate the difference in simulated outcomes

[18]Consider the deterministic case in which the plant, conditional on converting at some point prior to MATS (2015), decides when to convert. If operating costs and revenues for the coal versus natural gas unit are equal, the plant always decides to convert in the final period preceding MATS (e.g., $\beta \cdot \kappa < \kappa$). However, in the non-deterministic case, unobserved (to the researcher) shocks to fixed and/or operating costs may cause an earlier conversion. This is what is observed in the data, and generated by the model (via the logit-error).

[19]Controversy around the environmental impact of pipeline construction projects has lead to significant delays in the permitting process. Approval for most proposed projects now require extensive environmental impact analyses. Anecdotally, this process can take several years, though most spur projects are within two years for permitting and construction (https://www.bloomberg.com/news/articles/2016-07-21/want-to-build-a-gas-pipeline-grab-a-good-book-and-have-a-seat).

[20]These cumulative effects derive directly of the simulated conversion probabilities (the probability of being converted, conditional on not previously being converted). Specifically, I calculate this from the conditional probabilities as: $1 - \prod_{s=0}^{t}(1 - Pr(convert|\textit{not previously converted, observables})_s)$. These estimates are all relative to the base year in which the unit was defined as coal (2005).

induced by MATS versus a baseline, non-MATS simulated outcome. Though the estimates on conversions are very similar to those in Table A.13, we should expect the simulation results to produce smaller effects on conversions, as these calculations are mechanically different. Whereas Table A.13 presents estimates for the increase in likelihood of conversion for the treated group, Table A.16 calculates the increase in fraction of conversions from MATS across the entire sample.[21] Thus, this should be interpreted as the increased fraction of plants converted due to the implementation of MATS. As model results account for within plant spillovers of MATS from treated to control units, for comparison to Table A.13 estimates, it is arguably better to scale these plant-level effects by the fraction of plants with at least one treated unit (84.5%). Doing so produces estimates strikingly similar to reduced form estimates: approximately 10% for 2016.

In the next column, I calculate the *total* cost of MATS imposed on the 481 coal plants in my sample. Each annual estimate reports the difference in the coal plants' expected present value of future payoffs under MATS. As each surplus estimate is in expected present value, for consistency, I report the 2016 value in the final row. With a conservative estimate of a $1 million cost-per-mile of pipeline, MATS imposes a $26.8 billion dollar burden on coal plants. This translates into approximately $56 million per plant. As an alternative strategy for a plant to comply with MATS, anecdotal evidence suggests a scrubber system can cost between $50-160 million for a plant.[22] My estimates seem to be on par with this figure.

Next, I translate these conversions into environmental benefits. Each outcome is reported in 2015 U.S. dollars, and estimates the difference in pollutants produced under MATS, less the pollutants

[21] That is, Table A.13 reports $\mathbb{E}[converted|MATS = 1, treatment\ group = 1] - \mathbb{E}[converted|MATS = 0, treatment\ group = 1]$, where $MATS = 1$ indicates the universe where MATS takes place and $treatment\ group = 1$ indicates generating units with at least 25 megawatt capacity. The simulation results from Table A.16 report the effect on the entire sample: $\mathbb{E}[converted|MATS = 1] - \mathbb{E}[converted|MATS = 0]$. A simple scaling of these results by the fraction of treated units will not reproduce the reduced form results, as model results account for spillovers onto control units, within plant. However, an approximation may be made from the fraction of plants with at least one treated unit (84.5%), which yields results similar to those from Table A.13.

[22] This article presents a range of estimates, further stating that studies in 2000 and 2004 estimate the cost of dry scrubbers at about $50 million: http://gazette.com/cost-of-scrubbers-at-colorado-springs-power-plant-keeps-rising/article/1517307

produced in absence of MATS-induced conversions. Note that a natural gas conversion in one period continues to produce benefits in sequential periods in which it would otherwise be coal-operated. The social cost of carbon calculations (ΔSCC) fully incorporate the long-run effects of global greenhouse gas reductions coming from the conversions. Following full implementation of MATS, these savings amount to \$0.8-1 billion annually. I estimate the present value of future conversions by simulating these conversions for 540 years out of sample, at which point discounted values are sufficiently close to zero. This calculation begins in 2011, and therefore accounts for the aggregate impact of MATS. Present value calculations suggest a total of \$64 billion in greenhouse gas reductions, or about \$140 million per plant. I make similar calculations for local pollutants, sulfur dioxides (ΔSO_2) and nitrogen oxides (ΔNO_X), and report the sum of these external benefits (ΔExternal Benefits). The result is a meaningful \$83 billion external benefit from MATS, in terms of plant conversions.

These results suggest that MATS places substantial costs on coal plants, which are significantly smaller than the environmental benefits from the induced conversions. Note, however, that this analysis only captures reductions in emissions from three pollutants (though secondary pollutants are accounted for).[23] Furthermore, this analysis is restricted to coal-to-natural gas conversions. Therefore, this exercise is not meant to fully account for all costs and benefits of the MATS program, but rather, it's impact on conversions. In the next two sub-sections, I switch to the primary focus of this paper—the impact of pipeline infrastructure.

4.6.2 MATS, with Pipeline Subsidies

Next I examine how the MATS policy might produce even further environmental benefits when plants are granted subsidies for construction of lateral pipelines to the mainline natural gas network. These simulation results are presented in Table A.17 and illustrated graphically in Figure A.25. Note that I introduce the subsidies in the initial period, thus inducing conversions even prior

[23] Marginal damage estimates for SO2 and NOX also include damages from secondary pollutants produced from these emissions, such as particulate matter.

to MATS. Examining the effect of MATS in this scenario, we see that converted plants increase from 1% to 3% from 2010 to 2011, and up to 11% by the time of full implementation in 2015. This translates into a full percentage point increase in effectiveness from 2010 to 2011 compared to MATS alone. The 0.01 to 0.13 increase from pre- to post-MATS under pipeline subsidies is substantially larger than the 0.07 increase under MATS alone.

Producers also benefit significantly from this subsidy, as they no longer incur construction costs associated with connection to the pipeline network. These subsidies produce nearly a 10% reduction in the costs imposed on power plants, with a producer surplus impact of -$24.4 billion across all plants—about $51 million per plant. As for external benefits, MATS combined with pipeline subsidies produces a meaningful 23% higher overall benefit than that of MATS alone, including an additional $14.8 billion reduction in global greenhouse gases.

These benefits from pipeline subsidies generate significant costs, however. Given a price of $1 million per mile, simulation results suggest that coal-to-natural gas conversions under this scheme would cost the government $1.7 billion in present value, and about $100-$124 million a year for periods immediately following MATS implementation. These costs are so substantial because MATS already provides large incentives for these plants to convert to natural gas. Thus, this policy subsidizes plants who would have made the switch, even in absence of a subsidy. However, given a marginal benefit of adding subsidies of $19 billion (plus producer surplus benefits), which well exceeds the marginal cost, MATS and subsidized infrastructure should be considered complementary programs.

4.6.3 Subsidized Pipelines, Independent of MATS

In this section, I isolate the impact of the pipeline infrastructure by simulating plant conversions when faced with subsidized pipelines, but when under no external pressures on emissions in the form of MATS. The majority of the generating units converted in my sample have began operation prior to 1970. Thus, many of these units are likely old, inefficient coal units. In many circumstances, it may be optimal for a plant to convert an old, inefficient pre-existing coal unit to natural gas. This will often occur when the marginal plant anticipates higher variable costs of operating these units under coal than natural gas, but finds the fixed cost of converting too high to attain these benefits. This exercise examines how full subsidization of the infrastructure alone might encourage some of these conversions.

Though we should not expect this subsidy to be more effective than MATS mandates, results in Table A.18 suggest that plants respond significantly to these incentives by converting old coal-fired units to natural gas. Figure A.26 displays the plants' action graphically. Cumulative conversions from subsidized pipelines are over a third of that of conversions under MATS by 2015. For this simulation, I compute the total present value of external benefits beginning in 2006, as this is the relevant initial period in this scenario. From Table A.18, the estimates amount to \$10.1 billion in external benefits. These results show that coal plant pipeline connection costs are inhibiting \$7.9 billion in carbon dioxide, \$1.3 billion in sulfur dioxide, and \$923 million in nitrogen oxide reductions from taking place. Although, this policy scheme produces significant infrastructure spending: approximately \$602 million in present value.

We can use these simulation results to calculate the marginal reduction in damages per-mile of pipeline. As subsidy estimates assume a cost of \$1 million per-mile, Table A.18 demonstrates that this policy would produce between 28 and 33 miles of new pipeline per-year, within the time frame of the data. This implies immediate (contemporaneous) reductions in damages of between \$944 thousand and \$1.1 million per-mile of new pipeline each year. These estimates only apply for the subset of pipeline built within the time-frame of the data. Projecting forward all future projects, we can calculate the long-run marginal benefits for the entire pipeline infrastructure. Inherent in my model and this simulation, as plants are offered a subsidy for pipeline costs in each sequential period, at some time, T, all plants will convert. Thus, using the present value of benefits from all of these future conversions, we can calculate the marginal external benefits, per-mile. With external benefits of \$10.1 billion and total pipeline mileage needed to connect every plant of 3,831 miles, this produces an average marginal reduction in damages of \$2.6 per-mile. Note, however, that these calculations are limited to the scope of coal-to-natural gas conversions, and should therefore be considered a lower-bound. As only a limited number of conversions occur in each period, the present marginal value of subsidized pipeline is only \$157 thousand, suggesting substantial net benefits from this policy.

4.7 Conclusion

The EPA's new Mercury and Air Toxics Standards place significant costs on oil- and coal-fired power plants using electric generating units with capacities 25 megawatts or greater. Plants can comply with these standards by investing in costly emissions reduction equipment, or shutting down and opening a more efficient plant. Another option, which has become more common in recent years, is to convert these oil- and coal-fired boilers to natural gas.

This paper aims to estimate the role that the pipeline infrastructure plays in a power plant's decision to convert from coal to natural gas. Fully exploiting quasi-experimental variation in the MATS program, I employ a triple differences estimator in order to isolate the impact of pipeline proximity on the propensity to convert, under the lens of MATS. My estimates suggest that a 10% reduction in pipeline distance produces an additional 6.4% increase in likelihood of conversion for an EGU regulated under MATS.

This approach, however, only isolates the cross-sectional variation in the static pipeline infrastructure specifically under MATS. Therefore, I implement a dynamic model of the power plant's decision to convert, introducing structure into the mechanism in which the coal plant internalizes these pipeline costs. I fully exploit the MATS program within this framework as an as-good-as-random event which increases the cost of operating a standard coal electric generating unit, thus making pipeline proximity salient, where a natural gas conversion becomes a viable alternative.

This model isolates the impact of the current infrastructure on the plant's decision to convert, allowing me to examine alternative policies to encourage natural gas use over coal.

Primary estimates from counterfactual simulations suggest that pipeline subsidies can contribute to \$10.1 billion in external benefits from conversions, in terms of reduced carbon dioxide, sulfur dioxide, and nitrous oxide (and secondary sources of) pollution. Furthermore, these reductions in emissions come strictly from a reduction in the coal plants' fixed cost of conversion, which generates greater than one-third the effect on conversions as MATS mandates do. Additional calculations derive an implied marginal external benefit of \$2.6 million per-mile of pipeline. This analysis demonstrates that more favorable policy toward the natural gas infrastructure can provide significant long-term environmental benefits.

Burned natural gas emits about half the CO2 of coal. Though pipeline projects have their own short-run costs, understanding the benefits of these projects is crucial in approving new expansions. The results in this paper suggest significant benefits from the pipeline infrastructure. Therefore, this paper offers an additional policy lever that can not only encourage use of a natural gas, but can incentivize many old, inefficient coal plants to take on a cleaner source of energy.

5. SUMMARY AND CONCLUSIONS

This dissertation discusses various energy and environmental related externalities produced in market activities. The emphasis in each of these essays has primarily been focused on policy solutions to environmental outcomes. Particularly, in Chapter 2, I examined how consumer preferences in an arms race can stimulate the growth in size of the vehicle fleet. These types of preferences generate feedback loops in size, leading to costly externalities. I proposed one possible policy solution to internalize the externalities of an arms race—a Pigouvian tax on gasoline. Simulation results from a model of vehicle choice show with sufficiently high gas taxes, an arms race may be reversed. This has significant implications for Pigovian policy. Furthermore, I show that a static tax is not efficient when an arms race exists.

In Chapter 3 I examine consequences of perverse incentives produced by the National Flood Insurance Program (NFIP). Approximately 39% of the U.S. population resides on its shorelines, where shorelines only make up 10% of the land area. In the face of climate change and rising sea levels, it is important to consider potential adaptation efforts, and address these concentrations in flood-prone areas. In this essay, co-author Abigail Peralta and I find strong evidence that the NFIP is a major contributor to these concentrations in high risk areas. This implies that this policy has provided an incentive for people to take on more risk, leading to a significant increase in the cost of major floods. We estimate that NFIP contributed to a 6.6% increase in costs attributed to hurricane Katrina, and up to a 15% increase in costs associated with hurricane Harvey. These results suggest NFIP produces insurance that may be *too* affordable for households, and in order to reduce these costs, premiums must rise in flood-prone areas.

Finally, Chapter 4 examines infrastructure policy as an effective approach to encourage the use of clean by power plants. Given the majority of preexisting coal plants are greater than 50 years old, plant operators are used to producing electricity the same way they have for decades. However, new environmental regulations force coal plants to internalize a significant amout of the externalities that they produce. I examine the Environmental Protection Agency's (EPA's) Mercury and Air Toxics Standards (MATS) as one such policy. I show that when these policies are imposed on coal-fired generating units, the natural gas pipeline network becomes salient to plant operators. Operators will weight the costs of MATS with the fixed costs of connecting to the network and converting to natural gas, thereby avoiding these costs imposed on coal usage. These incentives may be an important component for regulators to consider for pipeline permit approval. Furthermore, I show that pipeline subsidies can bring about significant environmental long term benefits; a present value of up to $192 million in carbon reductions.

In-depth analyses of specific environmental externalities and potential policy solutions is especially important given growing concerns of climate change. Understanding the incentives of the entities involved is an important component. This dissertation has attempted to impact the policy debate on potential solutions to environmental problems, with strong causal analyses, using comprehensive empirical methods.

REFERENCES

Abadie, A., Diamond, A., and Hainmueller, J. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. Journal of the American Statistical Association, 105:493–505, 2010.

Agency, F. E. M. The 100-year base flood standard and the floodplain management executive order: A review. Prepared for the Office of Management and Budget., 1983.

Allcott, H. The welfare effects of misperceived product costs: Data and calibrations from the automobile market. American Economic Journal: Economic Policy, 5(3):30–66, 2013.

Allcott, H. and Knittel, C. Are consumers poorly informed about fuel economy? evidence from two experiments. American Economic Journal: Economic Policy, Forthcoming, 2018.

Allcott, H. and Wozny, N. Gasoline prices, fuel economy, and the energy paradox. Review of Economics and Statistics, 96(10):779–795, 2014.

Anderson, D. R. The national flood insurance program. problems and potential. The Journal of Risk and Insurance, 41(4):579–599, 1974.

Anderson, M. As the wind blows: The effects of long-term exposure to air pollution on mortality. Working Paper, 2016.

Anderson, M. L. and Auffhammer, M. Pounds that kill: The external costs of vehicle weight. Review of Economic Studies, 82(2):535–571, 2014.

Anderson, S. T., Kellogg, R., and Sallee, J. M. What do consumers believe about future gasoline prices? Journal of Environmental Economics and Management, 66(3):383–403, 2012.

Archsmith, J., Gillingham, K., Knittel, C. R., and Rapson, D. S. Attribute substitution in household vehicle portfolios. NBER Working Paper No. 23856, 2017.

Barreca, A., Clay, K., Deschenes, O., Greenstone, M., and Shapiro, J. S. Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. Journal of Political Economy, 124(1):105–159, 2016.

Bartik, A. W., Currie, J., Greenstone, M., and Knittel, C. R. The local economic and welfare consequences of hydraulic fracturing. National Bureau of Economic Research Working Paper 2360, 2017.

Bento, A., Gillingham, K., and Roth, K. The effect of fuel economy standards on vehicle weight dispersion and accident fatalities. NBER Working Paper No. 23340, 2017.

Berry, S., Levinsohn, J., and Pakes, A. Automobile prices in market equilibrium. Econometrica: Journal of the Econometric Society, pages 841–890, 1995.

Bertrand, M., Duflo, E., and Mullainathan, S. How much should we trust differences-in-differences estimates? Quarterly Journal of Economics, 119(1):249–275, 2004.

Bhat, C. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transporation Research B, 35:677–693, 2001.

Boustan, L. P., Kahn, M. E., and Rhode, P. W. Moving to higher ground: Migration response to natural disasters in the early twentieth century. American Economic Review, 102(3):238–244, 2012.

Boustan, L. P., Kahn, M. E., Rhode, P. W., and Yanguas, M. L. The effect of natural disasters on economic activity in us counties: A century of data. Technical report, National Bureau of Economic Research, 2017.

Bouwer, L. M., Crompton, R. P., Faust, E., Hoppe, P., and Jr., R. A. P. Confronting disaster losses. American Economic Journal: Economic Policy, 9(3):168–198, 2017.

BPA. Comparing pipes & wires. a joint study by the bonneville power administration and the northwest gas association. Available at http://www.chpcenternw.org (accessed August 1, 2009), 2004.

Brauner, E. V., Forchhammer, L., Moller, P., Simonsen, J., Glasius, M., Wahlin, P., Raaschou-Nielsen, O., and Loft1, S. Exposure to ultrafine particles from ambient air and oxidative stress-induced dna damage. Environmental Health Perspectives, 115(8):1177–82, 2007.

Brehm, P. Natural gas prices, electric generation investment, and greenhouse gas emissions. Working Paper, 2017.

Bresnahan, T. F. Competition and collusion in the american automobile industry: The 1955 price war. The Journal of Industrial Economics, XXXV(4):457–482, 1987.

Burby, R. J. Flood insurance and floodplain management: the us experience. Global Environmental Change Part B: Environmental Hazards, 3(3-4):111–122, 2001.

Busse, M., Knittel, C. R., and Zettelmeyer, F. Are consumers myopic? evidence from new and used car purchases. American Economic Review, 103(1):220–256, 2013.

Busse, M. R., Pope, D. G., Pope, J. C., and Silva-Risso, J. The psychological effect of weather on car purchases. Quarterly Journal of Economics, 131(1):371–414, 2014.

Busse, M. R., Knittel, C. R., Silva-Risso, J., and Zettelmeyer, F. Who is exposed to gas prices? how gasoline prices affect automobile manufacturers and dealerships. Quantitative Marketing and Economics, Forthcoming, 2016.

Cawley, J. and Philipson, T. An empirical examination of information barriers to trade in insurance. American Economic Review, 89(4):827–846, 1999.

Chay, K. and Greenstone, M. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. The Quarterly Journal of Economics, 118(3):1121–67, 2003.

Chay, K., Dobkin, C., and Greenstone, M. The clean air act of 1970 and adult mortality. Journal of Health Economics, 28(3):688–703, 2009.

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. Evidence on the impact of sustained exposure to air pollution on life expectancy from chinaâĂŹs huai river policy. Proceedings of the National Academy of Sciences, 110(32):12936–41, 2013.

Chivers, J. and Flores, N. E. Market failure in information: The national flood insurance program. Land Economics, 78(4):515–521, 2002.

Clay, K., Lewis, J., and Severnini, E. Canary in a coal mine: Infant mortality, property values, and tradeoffs associated with mid-20th century air pollution. NBER Working Paper No. 22155, 2016.

Clay, K., Jha, A., Muller, N. Z., and Walsh, R. The external costs of transporting petroleum products by pipelines and rail: Evidence from shipments of crude oil from north dakota. NBER

Working Paper No. w23852, 2017.

Conlin, M., O'Donoghue, T., and Vogelsang, T. J. Projection bias in catalog orders. American Economic Review, 97(4):1217–1249, 2007.

Crandall, R. W. and Graham, J. D. The effects of fuel economy standards on automobile safety. Journal of Law and Economics, 32:97–118, 1989.

Currie, J. and Neidell, M. Air pollution and infant health: What can we learn from californiaâĂŹs recent experience? The Quarterly Journal of Economics, 120(3):1003–30, 2005.

Currie, J., Neidell, M., and Schmieder, J. F. Air pollution and infant health: Lessons from new jersey. Journal of Risk and Uncertainty, 27(3):279–300, 2003.

Deryugina, T. The fiscal cost of hurricanes: Disaster aid versus social insurance. American Economic Journal: Economic Policy, 9(3):168–198, 2017.

Dionne, G., Michaud, P.-C., and Dahchour, M. Separating moral hazard from adverse selection and learning in automobile insurance: longitudinal evidence from france. Journal of the European Economic Association, 11(4):897–917, 2013.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. Selection on moral hazard in health insurance. American Economic Review, 103(1):178–219, 2013.

Evans, L. Causal influence of car mass and size on driver fatality risk. American Journal of Public Health, 91(7):1076–1081, 2001.

Fell, H. and Kaffine, D. T. The fall of coal: Joint impacts of fuel prices and renewables on generation and emissions. American Economic Journal: Economic Policy, 10(2):90–116, 2018.

Feyrer, J., Mansur, E. T., and Sacerdote, B. Geographic dispersion of economic shocks: Evidence from the fracking revolution. American Economic Review, 107(4):1313–34, 2017.

Furnham, A. A literature review of the anchoring effect. The Journal of Socio-Economics, 40(1): 35–42, 2011.

Gallagher, J. Learning about an infrequent event: Evidence from flood insurance take-up in the us. American Economic Journal: Applied Economics, 6(3):206–233, 2014.

Gregory, J. The impact of post-katrina rebuilding grants on the resettlement choices of new orleans homeowners. Working Paper, 2014.

Hausman, C. and Kellogg, R. Welfare and distributional implications of shale gas. Brookings Papers on Economic Activity, pages 71–125, 2015.

Hendel, I. and Nevo, A. Intertemporal substituion and storable products. Journal of the European Economic Association, 2(2-3):536–547, 2004.

Hensher, D. The valuation of commuter travel time savings for car drivers: Evaluating alternative model specifications. Transportation, 28:101–118, 2001.

Hotz, V. J. and Miller, R. A. Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies, 60(3):497–529, 1993.

Jacobsen, M. R. Fuel economy and safety: The influences of vehicle class and driver behavior. American Economic Journal: Applied Economics, 5(3):1–26, 2013.

Jha, A. and Muller, N. Z. The local environmental costs of coal procurement at u.s. power plants. NBER Working Paper No. 23417, 2017.

Joskow, P. L. Natural gas: From shortages to abundance in the united states. American Economic Review, 103(3):338–43, 2013.

Kahn, M. E. The death toll from natural disasters: The role of income, geography, and institutions. Review of Economics and Statistics, 87(2):271–284, 2005.

Kane, J. and Puentes, R. In flood-prone areas, a rising tide of population. Brookings, 2015.

Keane, M. and Stavrunova, O. Adverse selection, moral hazard and the demand for medigap insurance. Journal of Econometrics, 190(1):62–78, 2016.

Knittel, C. R., Metaxoglou, K., and Trindade, A. Natural gas prices and coal displacement: Evidence from electricity markets. NBER Working Paper No. 21627, 2015.

Kousky, C. and Shabman, L. Pricing flood insurance: how and why the nfip differs from a private insurance company. Resources for the Future Discussion Paper, 2014.

Kriesel, W. and Landry, C. Participation in the national flood insurance program: An empirical analysis for coastal properties. Journal of Risk and Insurance, 71(3):405–420, 2004.

Landry, C. E. and Jahan-Parvar, M. R. Flood insurance coverage in the coastal zone. Journal of

Risk and Insurance, 78(2):361–388, 2011. ISSN 1539-6975. doi: 10.1111/j.1539-6975.2010.01380.x. URL http://dx.doi.org/10.1111/j.1539-6975.2010.01380.x.

Larrick, R. and Soll, J. The mpg illusion. Science, 320(5883):1593–94, 2008.

Li, S. Traffic safety and vehicle choice: Quantifying the effects of the 'arms race' on american roads. Journal of Applied Econometrics, 27:34–62, 2012.

Lowenstein, G., O'Donoghue, T., and Rabin, M. Projection bias in predicting future utility. Quarterly Journal of Economics, 118:1209–1248, 2003.

Maantay, J. and Maroko, A. Mapping urban risk: Flood hazards, race, & environmental justice in new york. Applied Geography, 29(1):111–124, 2009.

Michel-Kerjan, E. Catastrophe economics: The national flood insurance program. Journal of Economic Perspectives, 24(4):399–422, 2010.

Michel-Kerjan, E. and Kousky, C. Come rain or shine: Evidence on flood insurance purchases in florida. The Journal of Risk and Insurance, 77(2):369–397, 2010.

Michel-Kerjan, E. and Kunreuther, H. Redesigning flood insurance. Science, 333(6041):408–409, 2011.

Morrissey, W. FemaâĂŹs flood hazard map modernization initiative, crs report for congress, 2006.

Muehlenbachs, L., Spiller, E., and Timmins, C. The housing market impacts of shale gas development. American Economic Review, 105(12):3633–59, 2015.

Muller, N. Z. and Mendelsohn, R. Measuring the damages of air pollution in the united states. Journal of Environmental Economics and Management, 54(1):1–14, 2007.

Muller, N. Z. and Mendelsohn, R. Efficient pollution regulation: Getting the prices right. American Economic Review, 99(5):1714–39, 2009.

Munizaga, M. and Alvarez-Daziano, R. Mixed logit vs. nested logit and probit models. Presented at 5th Tri-annual Invitational Choice Symposium Workshop: Hybrid Choice Models, Formulation and Practical Issues, 2001.

Neidell, M. Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. Journal of Health Economics, 23(6), 2004.

Nevo, A. Mergers with differentiated products: The case of the ready-to-eat cereal industry. The Rand Journal of Economics, 31(3):395–421, 2000.

Oliver, M. E. Economies of scale and scope in expansion of the u.s. natural gas pipeline network. Energy Economics, 52(B):265–276, 2015.

Petrin, A. and Train, K. A control function approach to endogeneity in consumer choice models. Journal of Marketing Research, XLVI, 2009.

Pierce, K. Soi migration data: A new approach. IRS Statistics of Income Bulletin, 2015.

Pralle, S. Drawing lines: Fema and the politics of mapping flood zones. Technical report, American Political Science Association Annual Conference, 2017.

Rappaport, J. and Sachs, J. D. The united states as a coastal nation. Journal of Economic Growth, 8(1):5–46, 2003.

Robinson, M. F. History of the 1% chance flood standard. Reducing flood losses: Is the, 1:2–8, 2004.

Rust, J. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. Econometrica, 55(5):999–1033, 1987.

Sallee, J. M., West, S. E., and Fan, W. Do consumers recognize the value of fuel economy? evidence from used car prices and gasoline price fluctuations. Journal of Public Economics, 135:61–73, 2016.

Slovic, P. The relative influence of probabilities and payoffs upon perceived risk of a gamble. Psychonomic Science, 9:223–224, 1967.

Smith, C. E. Natural gas pipeline profits, construction both up. Oil & Gas Journal, 114(9):84–102, 2016.

Solon, G., Haider, S. J., and Wooldridge, J. What are we weighting for? Journal of Human Resources, 50(2):301–16, 2015.

Spanier, J. and Maize, E. Quasi-random methods for estimating integrals using relatively small samples. SIAM Review, 36:18–44, 1991.

Spenkuch, J. L. Moral hazard and selection among the poor: Evidence from a randomized experiment. Journal of Health Economics, 31(1):72–85, 2012.

Strobl, E. The economic growth impact of hurricanes: evidence from us coastal counties. Review of Economics and Statistics, 93(2):575–589, 2011.

TMAC. Technical mapping advisory council future conditions risk assessment and modeling report, 2015.

Train, K. Halton sequences for mixed logit. Department of Economics, UCB, 2000.

Tversky, A. and Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 185 (4157):1124–1131, 1974.

Wang, J., Ryan, D., and Anthony, E. J. Reducing the greenhouse gas footprint of shale gas. Energy Policy, 30(12):8196–8199, 2011.

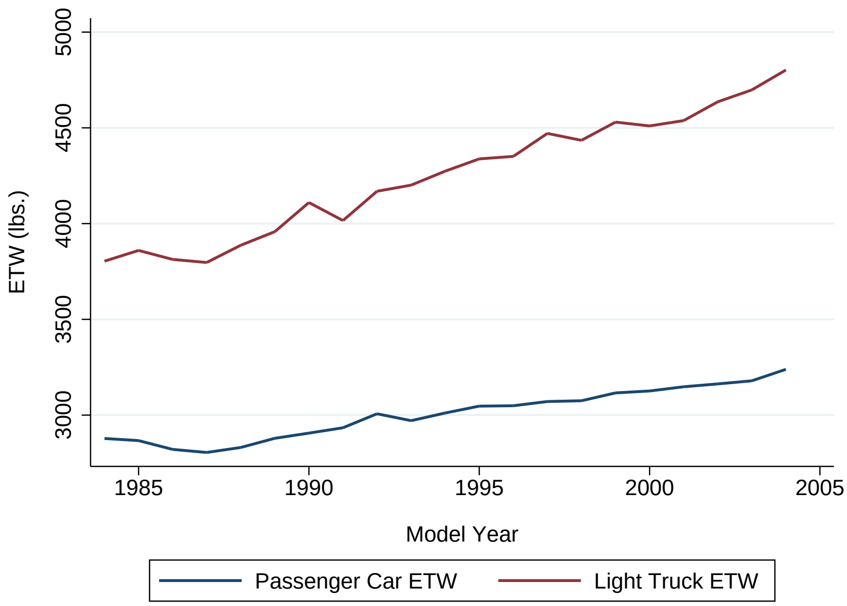
Weisburd, S. Identifying moral hazard in car insurance contracts. Review of Economics and Statistics, 97(2):301–313, 2015.

White, M. J. The ‘arms race’ on american roads: the effect of suvs and pickup trucks on traffic safety. Journal of Law and Economics, 47:333–356, 2004.

APPENDIX A

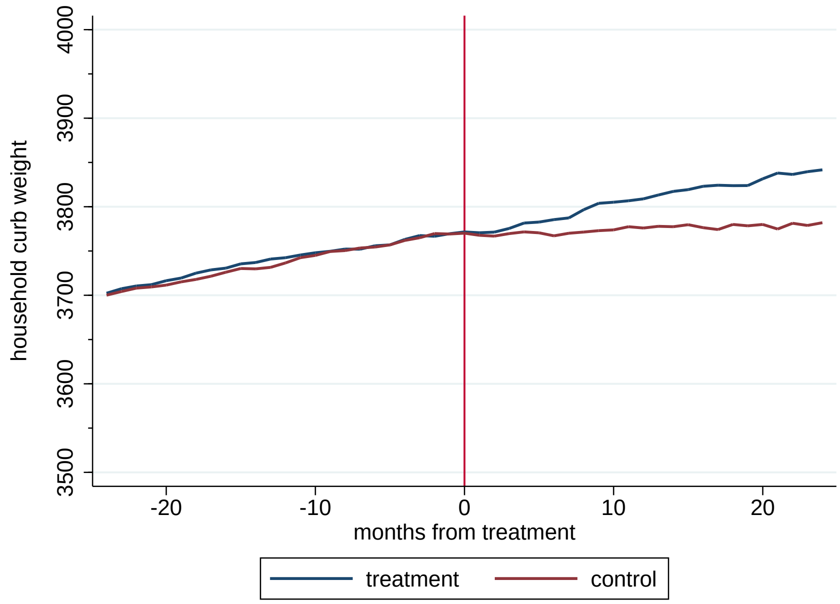
FIGURES AND TABLES

Figure A.1: U.S. Average Equivalent Test Weight (ETW)



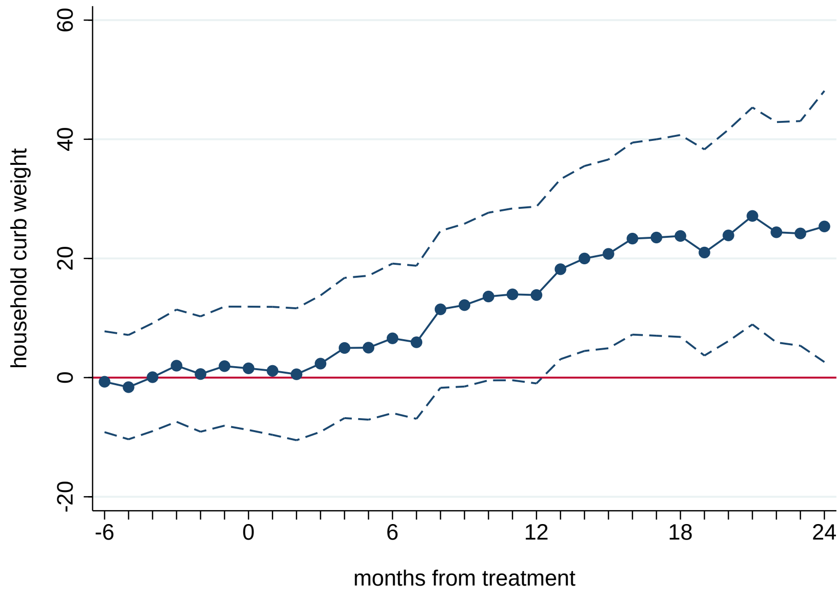
Note: This figure plots average Equivalent Test Weight (ETW) for passenger cars and light trucks. Data for this figure come from the National Highway Traffic Safety Administration (NHTSA).

Figure A.2: Effect of Neighbor's Fatality on Household Vehicle Size



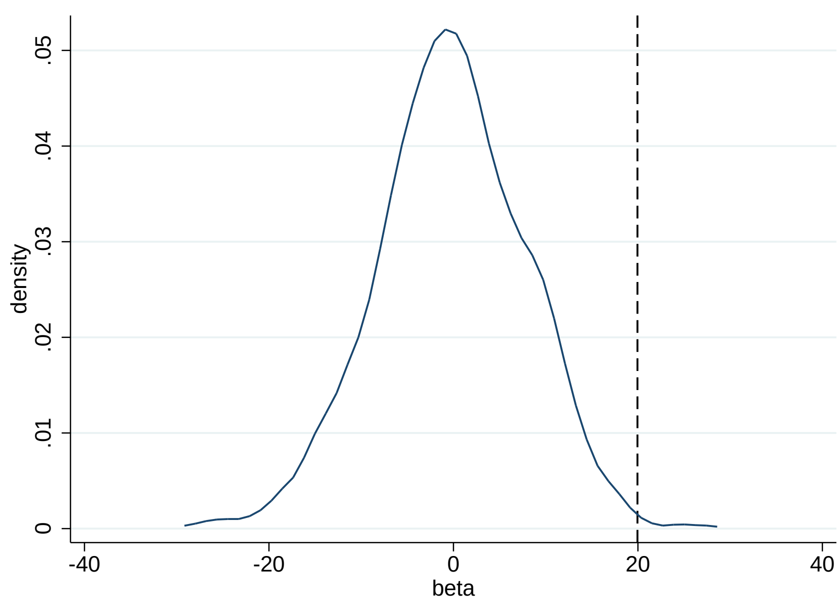
Note: This figure plots the average curb weight of the households' fleets, by month, for the treatment group and the control group. The horizontal axis plots the number of months before the nextdoor neighbors' vehicle accident occurred. The treatment group includes households neighboring someone who dies in an accident, whereas the control group includes households neighboring someone who survives an accident.

Figure A.3: Dynamic Estimates for Weight



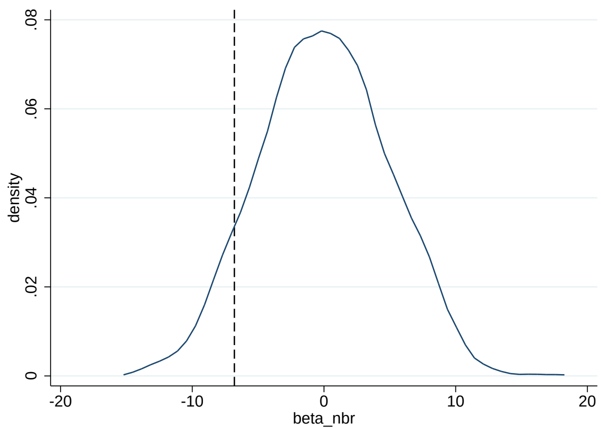
Note: This figure directly tests the parallel trends assumption of the difference-in-differences estimator. Each point represents an estimated coefficient on a lagged or lead treatment indicator. The lagged terms serve as pacebo periods, and test for divergence in the pre-treatment periods.

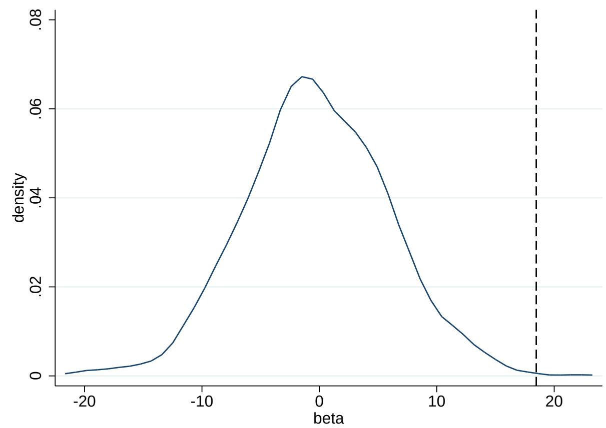
Figure A.4: Distribution of Placebo Estimates on Household Weight



Note: This is the resulting distribution of 1,000 simulated placebo estimates, generated by a series of randomly assigned placebo treatment periods and treatment groups. The vertical line corresponds to the estimated value of 19.95, with a p-value of 0.3%.

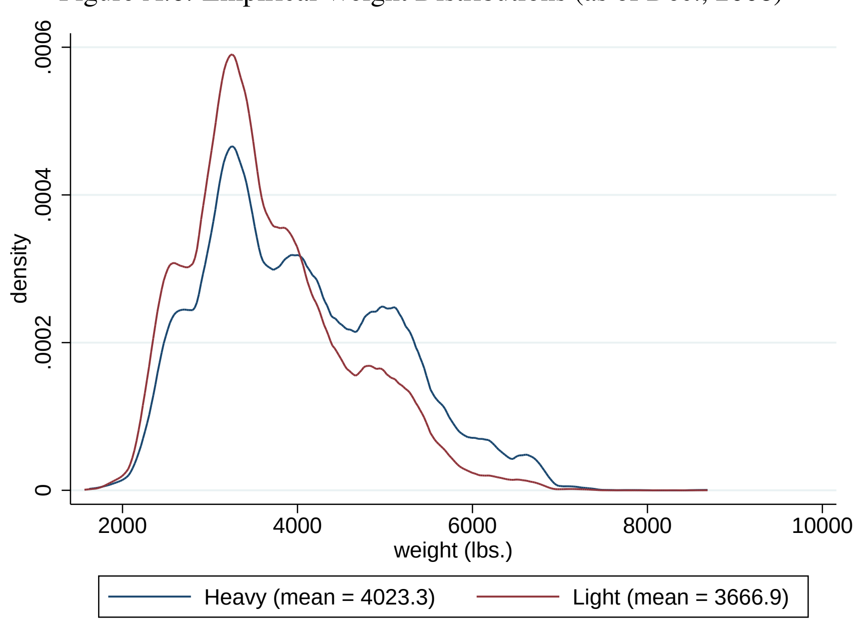
Figure A.5: Distribution of Placebo Estimates: Effect of Crash Vehicle Weights (Same Treatment Period, Random Neighbor)





Note: These are the resulting distributions of 1,000 simulated placebo estimates, generated by a series of randomly assigned placebo neighbors, while holding the true treatment period fixed. The vertical lines correspond to the estimated values of -6.80 and 18.47, for the neighbor and opposing vehicle, respectively. These estimates suggest one-tailed p-values of 0.082 and 0.001.

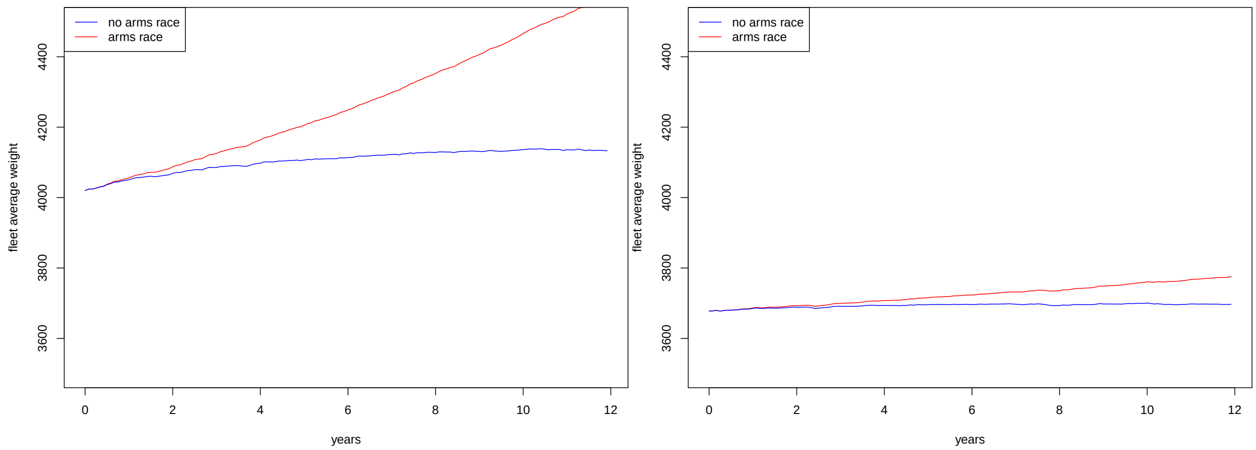
Figure A.6: Empirical Weight Distributions (as of Dec., 2008)



Note: Above plots the empirical weight distributions for the heavy (mean of 4,023 pounds) and light (mean of 3,667 pounds) counties used in the simulations, as of December 2008.

Figure A.7: Counterfactual Mean County Weights (fuel price = $1.665)

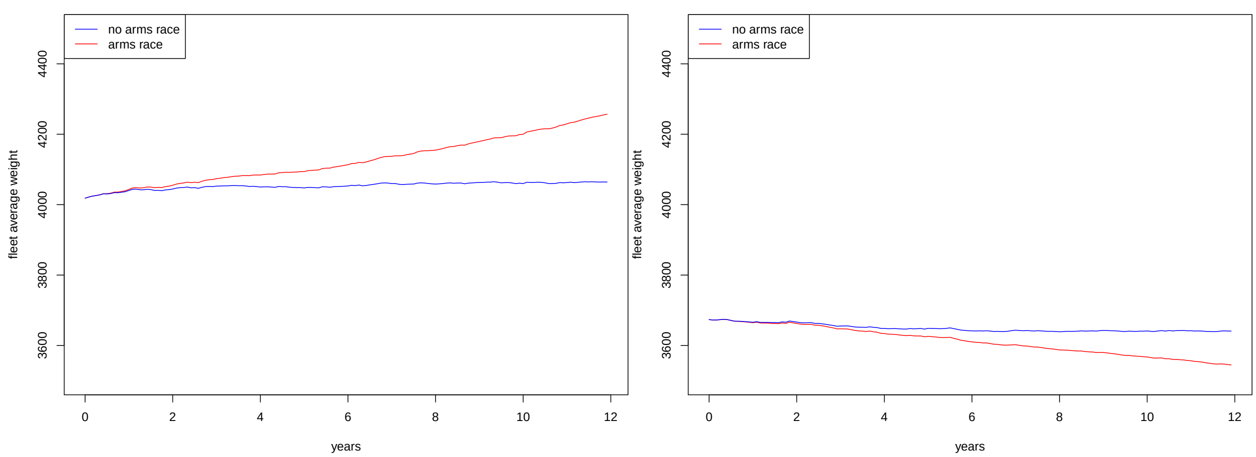
(a) Heavy County ($\Delta \bar{w}_T = 436$) (b) Light County ($\Delta \bar{w}_T = 78$)



Note: Simulated purchases are generated by month, assuming a monthly turnover rate of 2%, and a population size of 10,000. The initial weights in the population are drawn from the empirical distributions in December, 2008. The starting mean values at time 0 are equal to the means of these initial distributions. $\Delta \bar{w}_T$ indicates the difference between the arms race and no arms race groups at the end of the T periods.

Figure A.8: Counterfactual Mean County Weights (fuel price = $4)

(a) Heavy County ($\Delta \bar{w}_T = 193$) (b) Light County ($\Delta \bar{w}_T =$ -96)



Note: Simulated purchases are generated by month, assuming a monthly turnover rate of 2%, and a population size of 10,000. The initial weights in the population are drawn from the empirical distributions in December, 2008. The starting mean values at time 0 are equal to the means of these initial distributions. $\Delta \bar{w}_T$ indicates the difference between the arms race and no arms race groups at the end of the T periods.

Table A.1: Summary Statistics

Neighbors	1	2	3	5	10	ALL
Curb Weight	3,753.145	3,747.29	3,739.05	3,729.71	3,708.61	3,687.70
	(800.37)	(795.77)	(789.89)	(784.06)	(768.81)	(760.11)
Passenger Car*	0.64	0.65	0.65	0.65	0.66	0.67
	(0.48)	(0.48)	(0.48)	(0.48)	(0.47)	(0.47)
Airbag	0.81	0.81	0.80	0.81	0.81	0.81
	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)
MPG	18.99	18.99	19.00	19.03	19.10	19.16
	(3.81)	(3.80)	(3.81)	(3.80)	(3.80)	(3.80)
Purchases	2.16	2.17	2.16	2.15	2.15	2.13
	(2.24)	(2.20)	(2.17)	(2.15)	(2.14)	(2.13)
Households	7,923	15,297	22,081	34,039	56,159	78,220
Treated	0.58	0.58	0.57	0.57	0.56	0.56

*This refers to the household owning at least one passenger car, in contrast to all light trucks or SUVs.
Note: This table presents the sample mean for each variable, for each sample corresponding to the number of nextdoor neighbors included. Standard deviations are in parentheses.

*This refers to the household owning at least one passenger car, in contrast to all light trucks or SUVs.
Note: This table presents the sample mean for each variable, for each sample corresponding to the number of next-door neighbors included. Standard deviations are in parentheses.

Table A.2: Effect of Neighbor's Fatality on Household Vehicle Size (two nearest neighbors)

Outcome: Weight	(1)	(2)	(3)	(4)	(5)
neighbor died	32.16**	33.48**	34.50**	19.95**	16.37**
	(14.73)	(14.56)	(14.52)	(8.367)	(8.319)
Controls		Yes	Yes		
Month-Yr FE			Yes	Yes	Yes
HH FE				Yes	Yes
County Time Trend					Yes
N	964,824	964,824	964,824	964,824	964,824

* $p < 0.1$, ** $p < .05$, *** $p < .01$
Standard errors in parentheses, clustered at the street level. The outcome is household average vehicle size, and the samples includes the nearest two neighbors of the crash victims.

Table A.3: Effect of Neighbor's Fatality, by Neighbors

Neighbors	1	2	3	5	10
Outcome: Weight					
neighbor died	10.34	19.95**	15.21**	5.479	4.815
	(11.67)	(8.367)	(6.901)	(5.627)	(4.360)
N	495,278	964,824	1,397,654	2,159,147	3,590,583
Month $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Standard errors in parentheses, clustered at the street level. The outcome is household average vehicle size. Each column corresponds to a different sample of nearest neighbors of the crash victims (i.e., from 1 to 10 nearest neighbors).

Table A.4: Spillover Effects of Crash Vehicle Weights

	Neighbor Died		Household Weight				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
neighbor weight (1,000s)	-0.0713***	-0.0702***	-14.35*	-15.22*	-14.95*	-6.802	-7.395
	(0.00656)	(0.00651)	(8.161)	(8.050)	(8.025)	(5.062)	(4.991)
opposing weight (1,000s)	0.0581***	0.0577***	16.02	17.48*	16.85*	18.47***	17.32***
	(0.00605)	(0.00615)	(10.06)	(9.989)	(9.968)	(5.586)	(5.415)
Controls		Yes		Yes	Yes		
Yr-Month FE					Yes	Yes	Yes
HH FE						Yes	Yes
County Time Trend							Yes
N	3,780	3,780	414,817	414,817	414,817	414,817	414,817

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Standard errors in parentheses. In columns 1 and 2, standard errors are clustered at the level of the county of residence. These are the results of the first stage regression of the effect of neighbor's and opposing vehicle's weights on the probability of the neighbor dying. Columns 3-6 are clustered at the street level. The outcome is household average vehicle size for the two nearest neighbor sample.

Table A.5: Spillover Effects of Crash Vehicle Weights, by Neighbors

Neighbors	1	2	3	5	10
neighbor weight (1,000s)	-12.06*	-6.802	-5.669	-2.885	1.304
	(6.886)	(5.062)	(4.153)	(3.241)	(2.395)
opposing weight (1,000s)	17.38**	18.47***	10.84**	6.326*	6.962**
	(7.493)	(5.586)	(4.687)	(3.712)	(2.855)
Month $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes
N	212,762	414,817	598,491	928,146	1,551,218

* $p < 0.1$, ** $p < .05$, *** $p < .01$

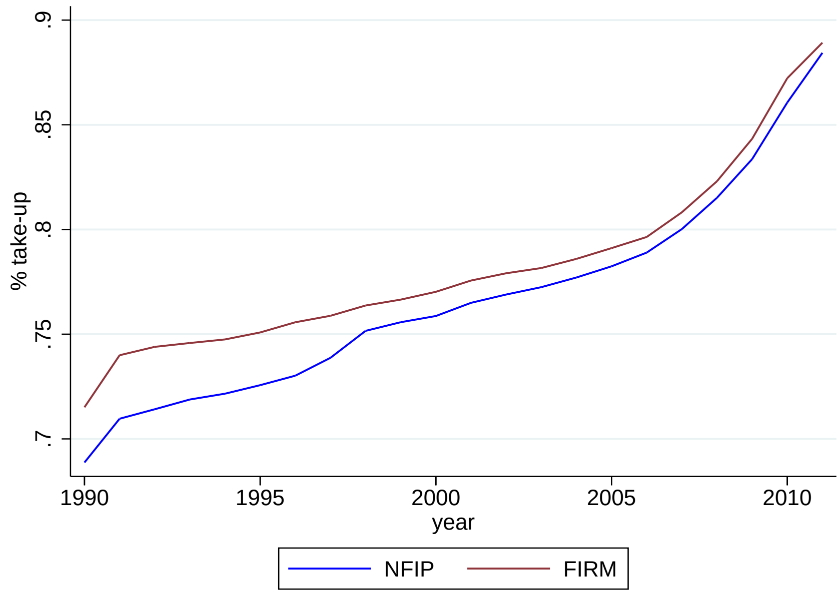
Standard errors in parentheses, clustered at the street level. The outcome is household average vehicle size. Each column corresponds to a different sample of nearest neighbors of the crash victims (i.e., from 1 to 10 nearest neighbors). Covariates of interest are the neighbor's and the opposing vehicle's weight.

Table A.6: Logit Demand Estimates

	Standard Logit	Price IV-Logit	Price & $\bar{w}$ IV-Logit		Random Coefficients	
price ($1,000s)	-0.0614	-0.0812	-0.0837	μ:	-2.3525	(mean = -0.0976)
	(0.0002)	(0.0002)	(0.0002)		(0.0152)	
				σ:	0.2253	
					(0.0023)	
$/mile	-12.7625	-6.4719	-5.4104		-4.6186	
	(0.0743)	(0.0955)	(0.0947)		(0.1472)	
horsepower	0.0044	0.0139	0.0155		0.0195	
	(0.0000)	(0.0001)	(0.0001)		(0.0014)	
weight (100 lbs.)	-0.0885	-0.1247	-0.7805		-0.7590	
	(0.0042)	(0.0042)	(0.0074)		(0.0107)	
$w \times \bar{w}$ (100^2 lbs.)	0.0042	0.0041	0.0221	μ:	0.0211	
	(0.0001)	(0.0001)	(0.0002)		(0.0019)	
				σ:	0.0018	
					(0.0007)	
WTP(w)/$\bar{w}$	$0.007	$0.005	$0.026		$0.023	
N	912,726	912,726	912,726		912,726	

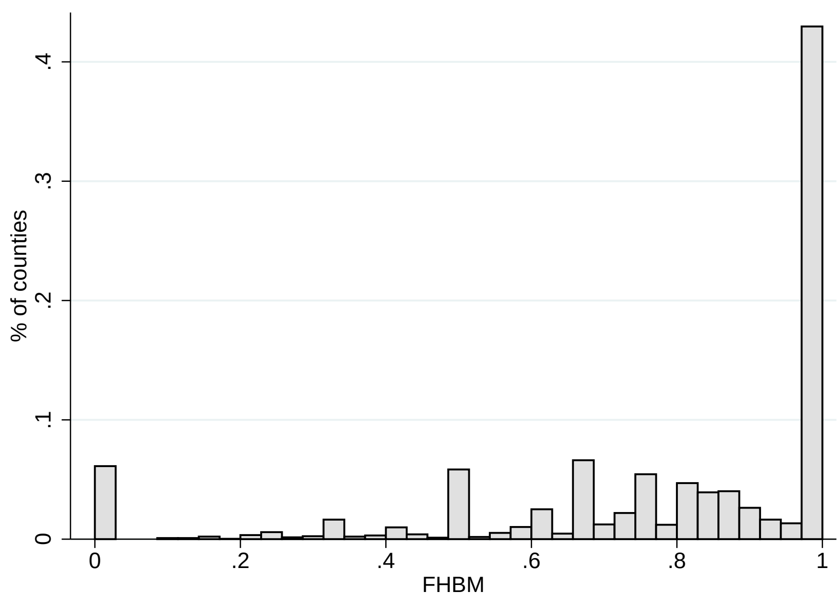
Above are the logit demand maximum likelihood estimates, asymptotic standard errors in parentheses. Column 1 presents standard logit estimates, including endogenous prices and mean weights. Column 2 presents control function estimates, instrumenting for price with BLP instruments. Column 3 presents control function estimates, instrumenting for price and mean weights. Column 4 presents simulated maximum likelihood estimates, instrumenting for price and mean weights using the control function approach, with random log-normal coefficients on price, and random normal coefficients on the interaction of weight and mean weights. The mean of the price coefficient is listed in parentheses. The row titled $WTP(w)/\bar{w}$ presents the estimates for the marginal willingness to pay for a pound, per additional mean weight pound.

Figure A.9: NFIP Enrollment and FIRM Assignment



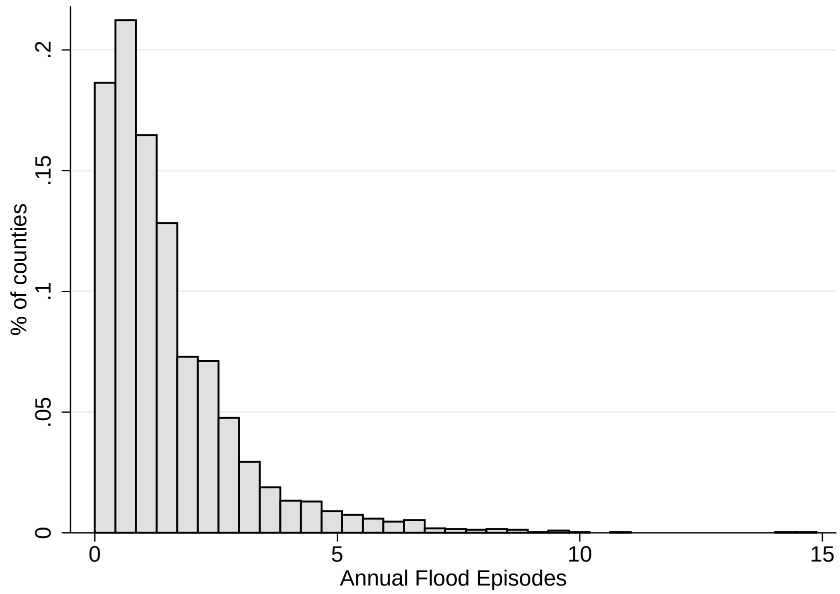
Note: This figure plots the fraction of counties enrolled in the National Flood Insurance Program (NFIP) and assigned a flood insurance rate map (FIRM) over time.

Figure A.10: Distribution of FHBM Counties



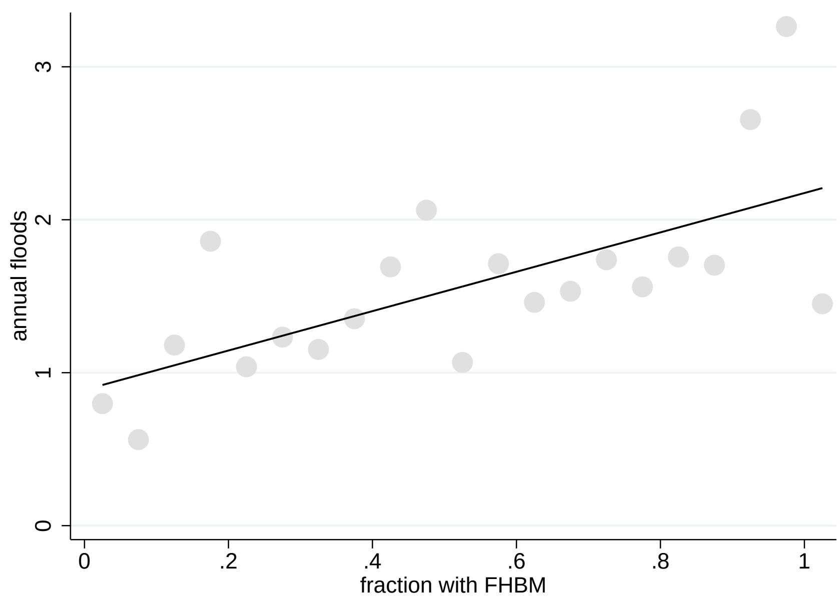
Note: This figure plots the distribution of fractions of a county (communities within a county) which FEMA has identified as flood-prone by publishing a Flood Hazard Boundary MAP (FHBM) for them. The vast majority of these assignments occurred in the 1970s.

Figure A.11: Distribution of Flood Risk Across Counties



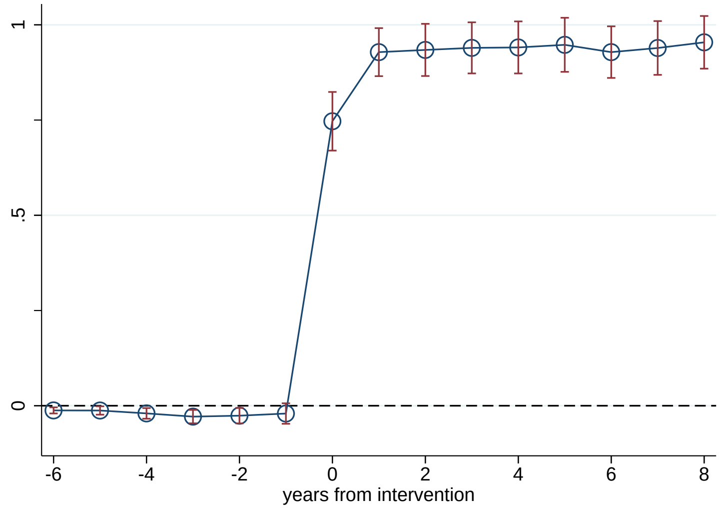
Note: This figure plots the distribution of county flood risk, defined as historical annual flood episodes (reported by the National Oceanic and Atmospheric Administration).

Figure A.12: Relationship Between FHBM and Flood Risk



Note: This figure presents a bin scatter plot of the fraction of a county with a Flood Hazard Boundary Map (FHBM) (on horizontal axis) and historical annual flood episodes (reported by the National Oceanic and Atmospheric Administration), with a bin size of 0.05 FHBM. The fitted line has a slope of 0.73 with a t-stat of 8.47, clustered by county.

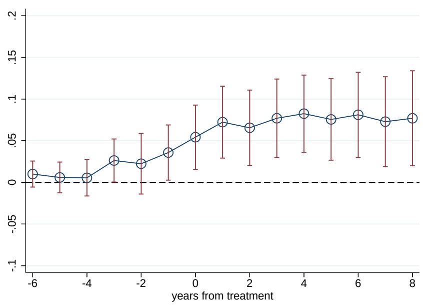
Figure A.13: First Stage: The Effect of FEMA Intervention on NFIP Enrollment



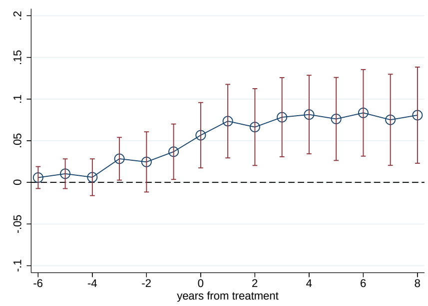
Note: This figure presents they dynamic coefficients from our first stage regression of the effect of post-FIRM FHBM intervention by FEMA—which granted affected areas with a 1 year grace period before sanctions were imposed—on actual NFIP enrollment. 95 percent confidence interval bars are presented. Standard errors clustered by county.

Figure A.14: Event Study Specification of Impact of (Endogenous) NFIP: Naïve Specification

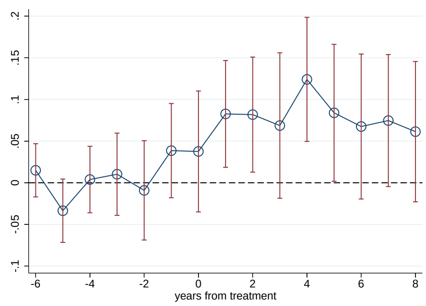
(a) Effect of NFIP on (log) Population



(b) Effect of NFIP on (log) Non-Migrants



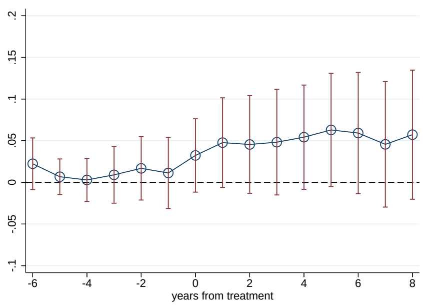
(c) Effect of NFIP on (log) Migrants



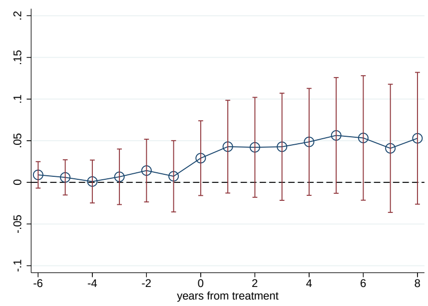
Note: This figure plots the coefficients from a naïve version of Equation 3.4, using lagged and leading National Flood Insurance (NFIP) enrollment indicators directly as the explanatory variables of interest. The final lagged term includes the entire post periods, thus, all coefficients are defined relative to the pre-NFIP period. Estimates are plotted for log-population (Panel a), log-non-migrants (Panel b), and log-migrants (Panel c) outcomes. 95% confidence intervals are presented. These results simply serve the purpose of illustrating the pre-divergence of treatment when directly using endogenous NFIP enrollment.

Figure A.15: Event Study Specification of Intent to Treat of NFIP

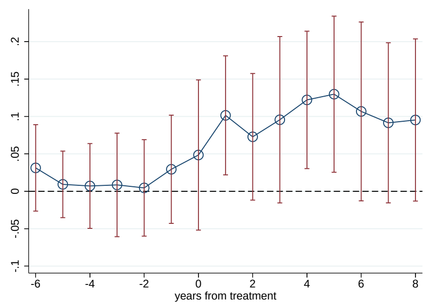
(a) Effect of NFIP on (log) Population



(b) Effect of NFIP on (log) Non-Migrants



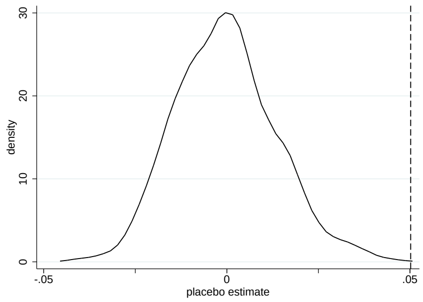
(c) Effect of NFIP on (log) Migrants



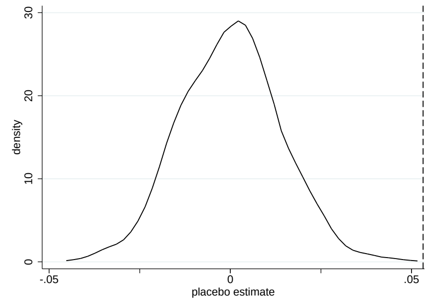
Note: This figure plots the reduced-form coefficients from Equation 3.4, using lagged and leading Flood Insurance Rate Map terms in our instrument. The final lagged term includes the entire post periods, thus, all coefficients are defined relative to the pre-FIRM period. Estimates are plotted for log-population (Panel a), log-non-migrants (Panel b), and log-migrants (Panel c) outcomes. 95% confidence intervals are presented.

Figure A.16: Placebo Estimates

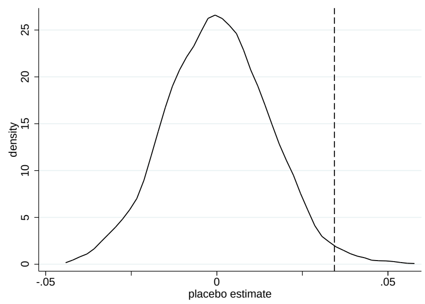
(a) Effect of NFIP on (log) Population



(b) Effect of NFIP on (log) Non-Migrants



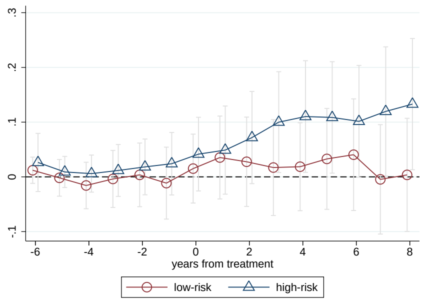
(c) Effect of NFIP on (log) Migrants



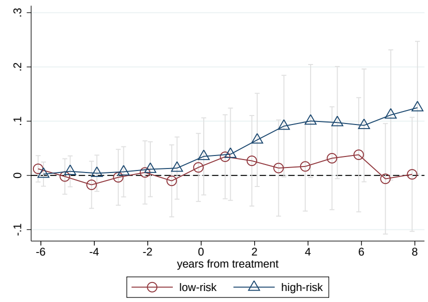
Note: Above are the resulting distributions of 1,000 placebo estimates of Equation 3.2 for each migration outcome. The dashed lines mark the corresponding point estimates from Table A.9. The implied p-values are 0.000, 0.000, and 0.015 for Panels a, b, and c, respectively.

Figure A.17: Heterogeneous Effects of NFIP, by Flood Risk

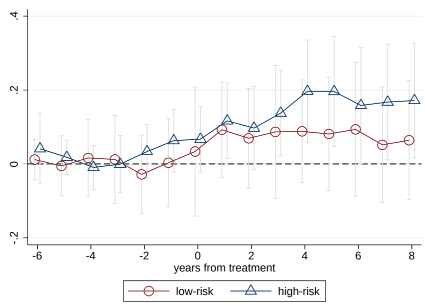
(a) Effect of NFIP on (log) Population



(b) Effect of NFIP on (log) Non-Migrants



(c) Effect of NFIP on (log) Migrants



Note: This figure plots the reduced-form coefficients from Equation 3.4, using lagged and leading Flood Insurance Rate Map terms in our instrument, from two separate samples—below (low-risk) and above (high-risk) median risk defined by annual historical flood episodes. The final lagged term includes the entire post periods, thus, all coefficients are defined relative to the pre-FIRM period. Estimates are plotted for log-population (Panel a), log-non-migrants (Panel b), and log-migrants (Panel c) outcomes. 95% confidence intervals are presented.

Table A.7: County Characteristics

NFIP	0.926
	(0.224)
FHBM	0.780
	(0.280)
Total Exemptions	71194.5
	(221581.6)
Non-Migrant Tax Exemptions	67151.2
	(212434.4)
Migrant Tax Exemptions	4042.0
	(10438.7)
Annual Flood Episodes	1.50
	(1.45)
Water-Related Decleared Disasters	0.143
	(0.450)
Per-Capita Income	26932.0
	(10731.3)
Unemployment Rate	6.484
	(3.290)

Note: Sample means and standard deviations (in parentheses) are presented. NFIP represents the proportion of communities ultimately enrolled in the National Flood Insurance Program. FHBM represents the proportion of communities assigned a flood hazard boundary map. We use number of tax exemptions as a proxy for population, where non-migrants are defined as returns filed in the same county in back to back years, and migrants refer to filings in a different county from one year to the next.

Table A.8: First Stage: NFIP enrollment on FIRM and FHBM assignment

Post-NFIP	(1)	(2)	(3)
Post-FIRM × FHBM	0.950***	0.950***	0.950***
	(0.0163)	(0.0163)	(0.0163)
County FE	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes
Controls		Yes	Yes
Declared Disaster Controls			Yes
County Time Trend			
N	64472	64472	64472

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Note: Above are the first stage estimates which regress National Flood Insurance Enrollment (NFIP) enrollment on flood hazard boundary map (FHBM) and flood insurance rate map (FIRM) assignment. In column 2, we control for county-level demographic variables, including per-capita income, unemployment rate, and job counts. Column 3 includes county-level water-related declared natural disasters. The estimates are not sensitive to the inclusion of these controls. Standard errors in parentheses are clustered on county. Note that 99% of counties with an FHBM ultimately receive a FIRM.

Table A.9: Effect of Flood Insurance on Migration (Reduced Form)

Migration Outcome	(1)	(2)	(3)	(4)	(5)
Panel A: Log- Population					
Post-FIRM × FHBM	0.0502***	0.0480***	0.0496***	0.0497***	0.0267*
	(0.0173)	(0.0163)	(0.0183)	(0.0183)	(0.0140)
FIRM Lead × FHBM			0.00747	0.00744	0.00302
			(0.0131)	(0.0131)	(0.00998)
Panel B: Log- Non-Migrants					
Post-FIRM × FHBM	0.0532***	0.0507***	0.0530***	0.0530***	0.0303**
	(0.0177)	(0.0166)	(0.0188)	(0.0188)	(0.0141)
FIRM Lead × FHBM			0.0105	0.0105	0.00644
			(0.0134)	(0.0134)	(0.0101)
Panel C: Log- Migrants					
Post-FIRM × FHBM	0.0344*	0.0343*	0.0313	0.0315	0.0107
	(0.0188)	(0.0182)	(0.0204)	(0.0204)	(0.0226)
FIRM Lead × FHBM			-0.0138	-0.0138	-0.0183
			(0.0177)	(0.0177)	(0.0178)
N	64472	64472	64472	64472	64472
County FE	Yes	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes	Yes
Declared Disaster Controls				Yes	Yes
County Time Trend					Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Note: Above are the OLS estimates of reduced-form Equation 3.2 of flood hazard boundary map (FHBM) and flood insurance rate map (FIRM) assignment on outcomes log-population (Panel a), log-non-migrants (Panel b), and log-migrants (Panel c). County-level demographic controls in columns 2-5 include per-capita income, unemployment rate, and job counts. County-level water-related declared natural disasters are included in columns 4-5. Leading terms of treatment are included in columns 3-5 as a falsification, and county-specific time trends are included in column 5. Standard errors in parentheses are clustered on county. Note that 99% of counties with an FHBM ultimately receive a FIRM.

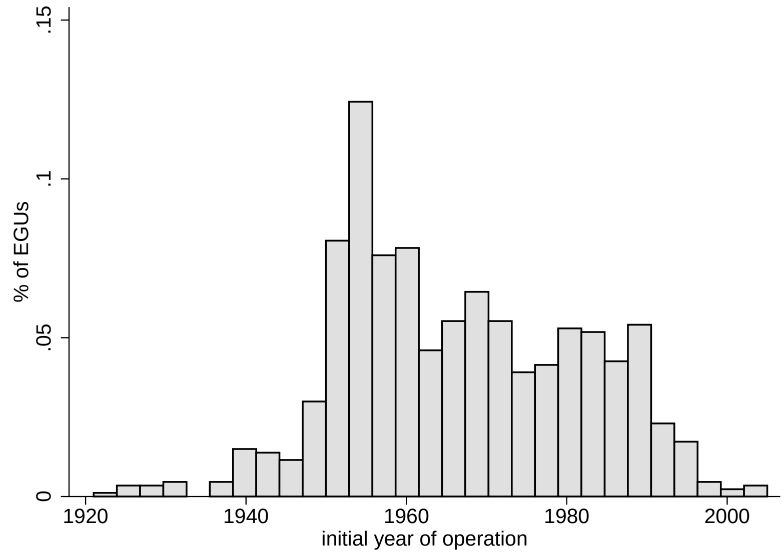
Table A.10: Heterogeneous Effects of NFIP on Migration, by Flood Risk

	Migration Outcome			
	(1)	(2)	(3)	(4)
Panel A: Log- Population				
Post-FIRM × FHBM	-0.0696***	0.00834	0.00158	0.00161
	(0.0266)	(0.0256)	(0.0248)	(0.0248)
Annual Floods × Post-FIRM × FHBM	0.0863***	0.0272*	0.0331**	0.0331**
	(0.0169)	(0.0151)	(0.0141)	(0.0141)
Panel B: Log- Non-Migrants				
Post-FIRM × FHBM	-0.0822***	0.00639	-0.00111	-0.00108
	(0.0311)	(0.0269)	(0.0261)	(0.0261)
Annual Floods × Post-FIRM × FHBM	0.0980***	0.0307*	0.0369**	0.0369**
	(0.0213)	(0.0163)	(0.0154)	(0.0154)
Panel C: Log- Migrants				
Post-FIRM × FHBM	-0.0265	0.0103	0.00963	0.00971
	(0.0268)	(0.0271)	(0.0264)	(0.0264)
Annual Floods × Post-FIRM × FHBM	0.0401**	0.0123	0.0158	0.0159
	(0.0157)	(0.0156)	(0.0149)	(0.0149)
N	67559	67559	67559	67559
County FE	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes
Year X Floods		Yes	Yes	Yes
Controls			Yes	Yes
Declared Disaster Controls				Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

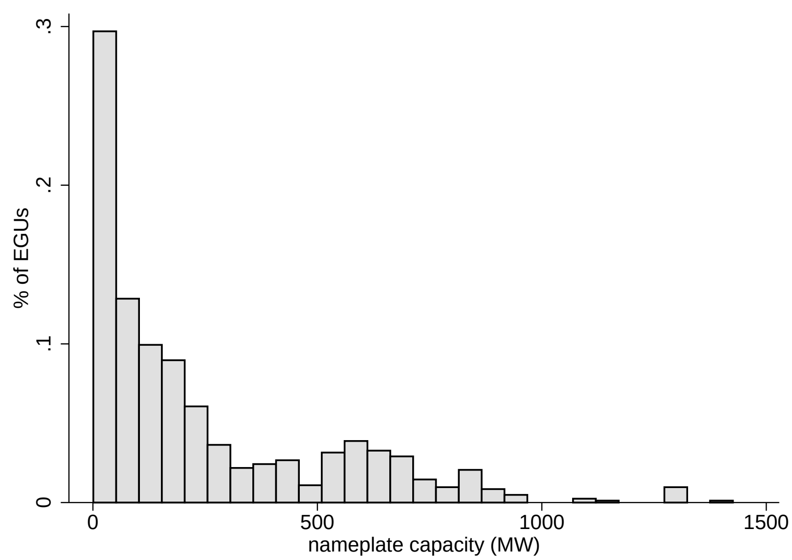
Note: Above are OLS estimates of our triple-diff specification in Equation 3.3 with outcomes log-population (Panel a), log-non-migrants (Panel b), and log-migrants (Panel c). Coefficients of interest are on the additional impact of NFIP (in reduced-form) from one additional flood per-year. Standard errors in parentheses are clustered on county.

Figure A.18: Coal Generating Unit Initial Year of Operation



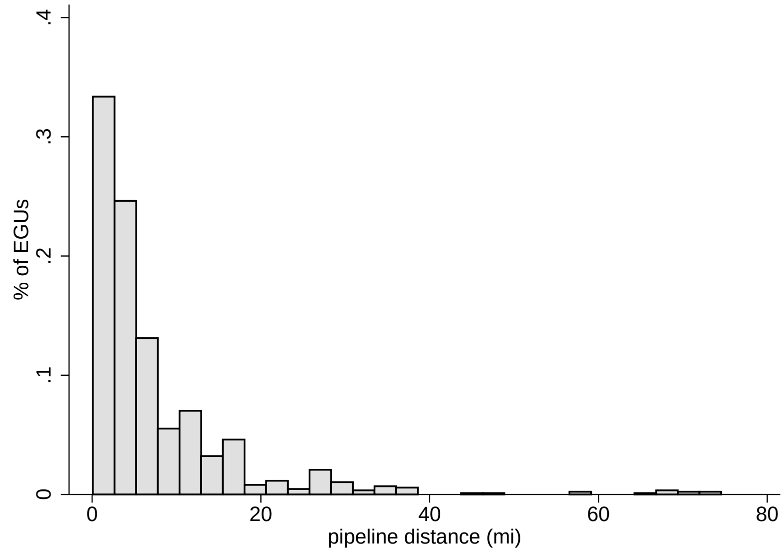
Note: Above plots the distribution of electric generating unit (EGU) initial operating year.

Figure A.19: Nameplate Capacity (MW) Distribution for Original Coal-Fired Units



Note: Above plots the distribution of electric generating unit (EGU) capacity. The vertical line notes 25 megawatts, where units greater or equal to 25 MW are subject to the Mercury and Air Toxics Standards (MATS).

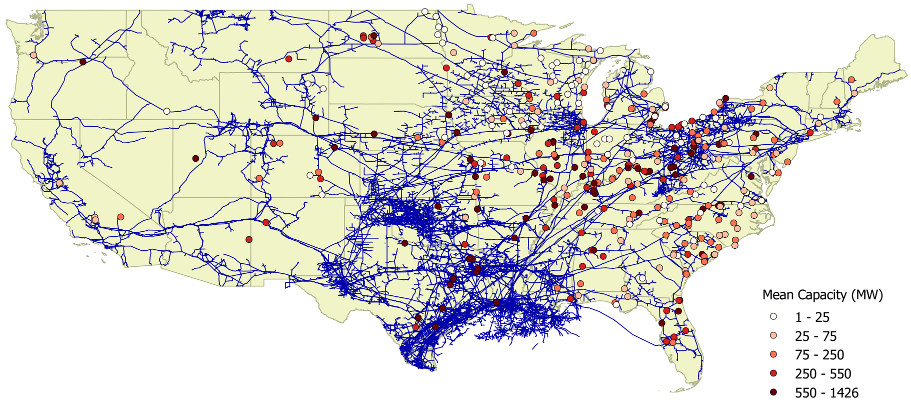
Figure A.20: Pipeline Distance



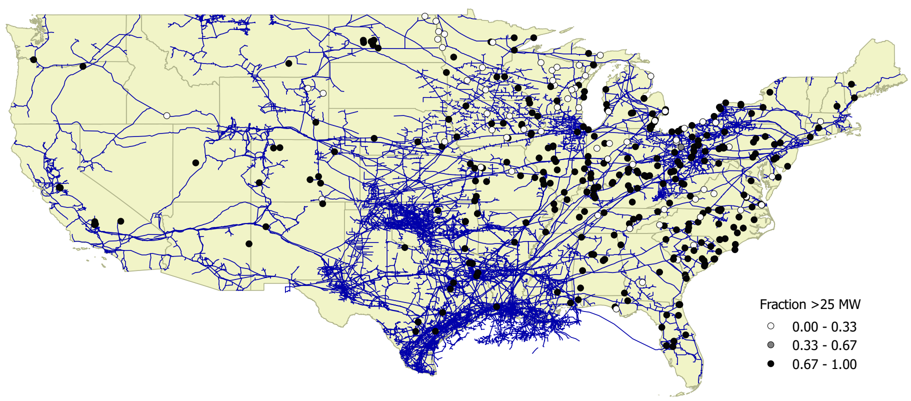
Note: Above plots the distribution of coal plant distances to nearest pipeline node, in miles.

Figure A.21: Power Plants Under MATS and the Natural Gas Pipeline Infrastructure

(a) Coal Plants' Mean EGU Capacity (MW)

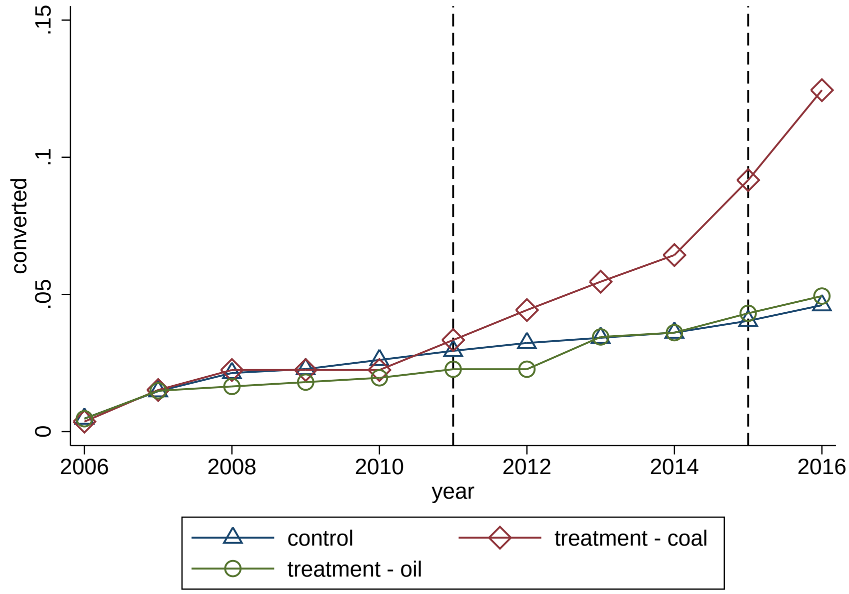


(b) Coal Plants Under MATS



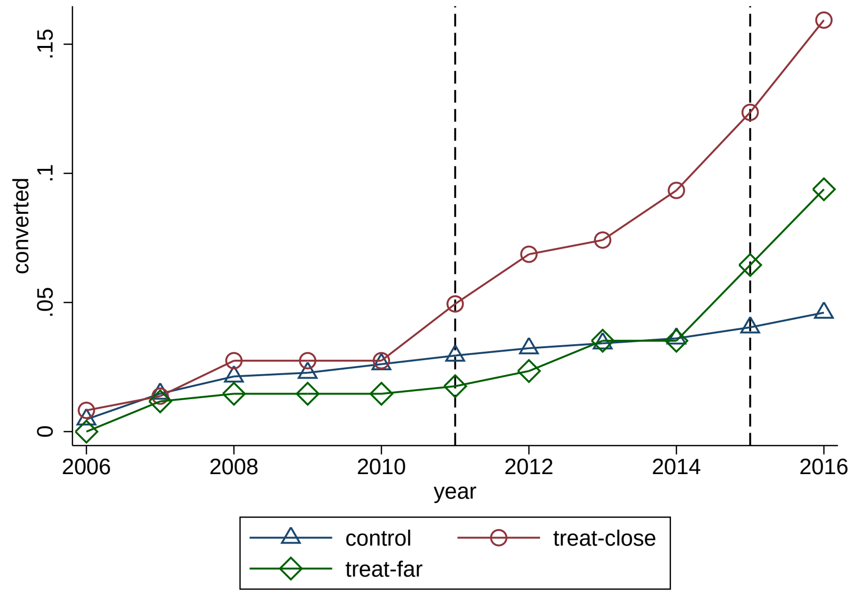
Note: Panel a presents variation in average plant-level capacity levels, whereas Panel b presents plants' fraction of units greater than 25 megawatts (or subject to MATS).

Figure A.22: Effect of MATS on EGU Conversions



Note: This figure plots the fraction of converted units over time for the control group (units < 25 MW), and coal and oil units in the MATS-treated group (units ≥ 25 MW). The vertical dashed lines represent years relevant to MATS: 2011, which was the announcement date, and 2015, which was the date of compliance.

Figure A.23: Effect of MATS on Coal EGU Conversions, by Pipeline Proximity

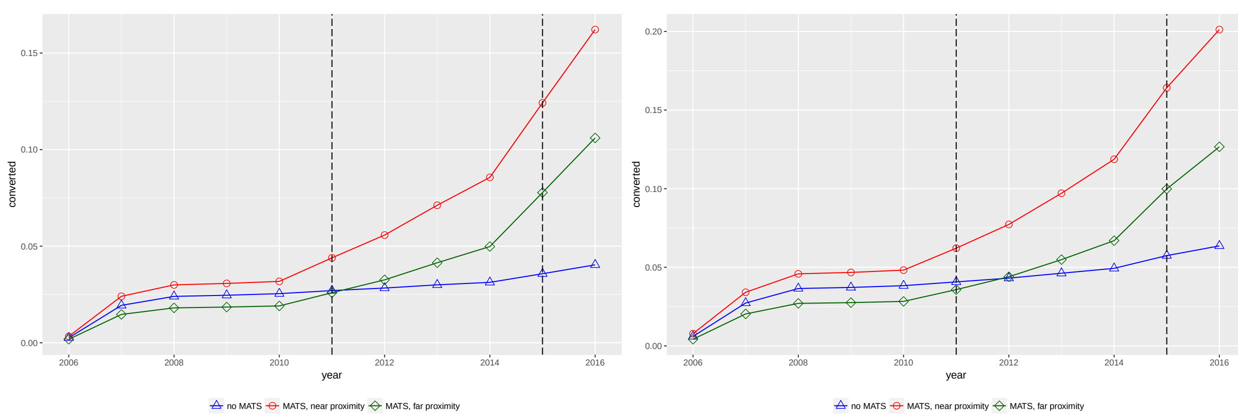


Note: This figure plots the fraction of converted units over time for the control group (units < 25 MW), and MATS-treated coal units (units ≥ 25 MW) below (close) and above (far) median proximity from nearest pipeline node. The median distance is 4.7 miles. The vertical dashed lines represent years relevant to MATS: 2011, which was the announcement date, and 2015, which was the date of compliance.

Figure A.24: Model Fitted Values of Below/Above Median Distance

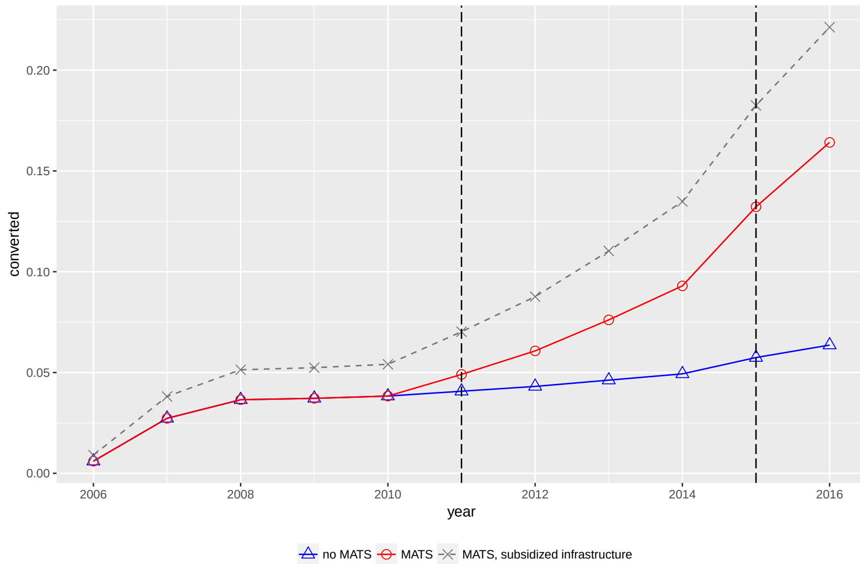
(a) MLE Estimates

(b) Weighted-MLE Estimates



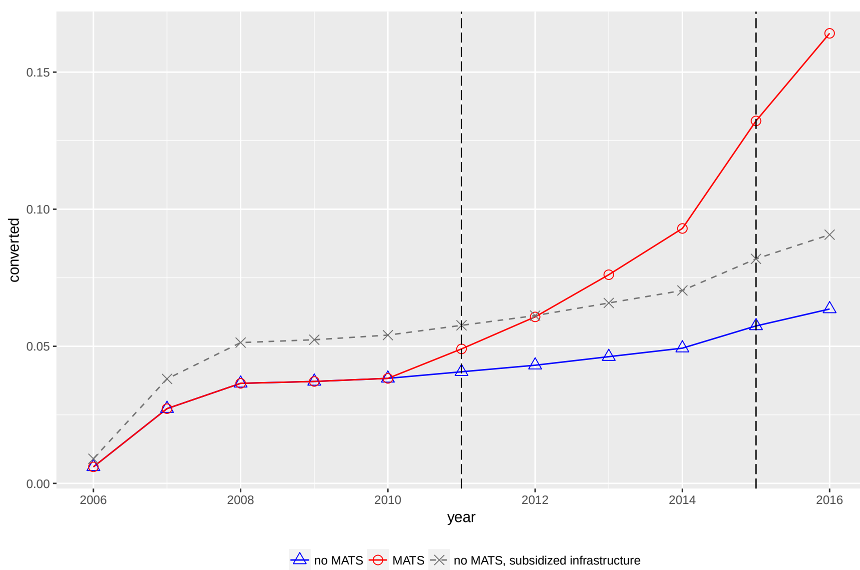
Note: This figure plots the *model-predicted* fraction of converted units over time for the control group (units < 25 MW), and MATS-treated coal units (units ≥ 25 MW) below (close) and above (far) median proximity from nearest pipeline node. This is analogous to the sample moments in Figure A.23. The vertical dashed lines represent years relevant to MATS: 2011, which was the announcement date, and 2015, which was the date of compliance. Panel a come from unweighted maximum likelihood estimates, whereas Panel b come from weighted maximum likelihood estimates, where I weight each likelihood by the inverse number of units in a plant's fleet.

Figure A.25: Policy Experiment: MATS + Fully Subsidized Pipelines



Note: This figure plots model predictions for the control group (units $<$ 25 MW), MATS-treated coal units (units $\geq$ 25 MW), and a counterfactual scenario of MATS *and* subsidized lateral pipelines to power plants. The vertical dashed lines represent years relevant to MATS: 2011, which was the announcement date, and 2015, which was the date of compliance.

Figure A.26: Policy Experiment: No MATS + Fully Subsidized Pipelines



Note: This figure plots model predictions for the control group (units $<$ 25 MW), MATS-treated coal units (units $\geq$ 25 MW), and a counterfactual scenario of subsidized lateral pipelines to power plants, with no units subject to MATS. The vertical dashed lines represent years relevant to MATS: 2011, which was the announcement date, and 2015, which was the date of compliance.

Table A.11: Summary Statistics

	no-MATS	MATS	All
Coal EGUs	164	756	1106
Oil EGUs	1970	942	2726
Converted	0.0461	0.0917	0.0648
	(0.210)	(0.289)	(0.246)
Nameplate Capacity (MW)	4.115	255.6	93.71
	(5.675)	(266.1)	(199.4)
Pipeline Distance (mi)	11.01	9.250	10.29
	(16.06)	(12.88)	(14.86)

Prime Mover	Combined-Cyle Combustion Turbine	Combustion (Gas) Turbine	Internal Combustion Engine	Steam Turbine
Percent	0.28	7.65	59.82	32.21
Converted	0.125	0.023	0.012	0.178

Note: Above are the sample statistics of key variables used in this paper. The first two rows present the number of coal and oil electric generating units (EGUs) observed in the data, split by whether they are regulated under the Mercury and Air Toxics Standards or not. The final column presents statistics for the full sample. Means, and standard deviations in parentheses, are presented for an indicator for whether the unit is converted in the time-frame of the data (between zero and one), its nameplate capacity in megawatts, and distance (in miles) from the pipeline network. The last two columns present the percent (between zero and one-hundred), and the mean converted (between zero and one) of each prime mover observed in the data.

Table A.12: Effect of Pipelines on Conversions Under MATS (OLS)

Outcome: Converted	(1)	(2)	(3)	(4)	(5)
MATS-2011	0.00608	0.00608	0.0116	0.0132	0.00831
	(0.00812)	(0.00812)	(0.0116)	(0.0102)	(0.0110)
MATS-2012	0.0160	0.0160	0.0252*	0.0268**	0.0215
	(0.00972)	(0.00972)	(0.0137)	(0.0125)	(0.0133)
MATS-2013	0.0226**	0.0226**	0.0315**	0.0331**	0.0275*
	(0.0109)	(0.0109)	(0.0150)	(0.0139)	(0.0146)
MATS-2014	0.0306**	0.0306**	0.0432***	0.0448***	0.0391**
	(0.0120)	(0.0120)	(0.0164)	(0.0155)	(0.0161)
MATS-2015	0.0561***	0.0561***	0.0706***	0.0722***	0.0663***
	(0.0149)	(0.0149)	(0.0193)	(0.0185)	(0.0190)
MATS-2016	0.0831***	0.0831***	0.102***	0.103***	0.0973***
	(0.0183)	(0.0183)	(0.0222)	(0.0216)	(0.0221)
MATS-2011 $\times$ distance (100 mi)			-0.0675	-0.0870***	-0.0423
			(0.0510)	(0.0333)	(0.0372)
MATS-2012 $\times$ distance (100 mi)			-0.112*	-0.131***	-0.0835*
			(0.0612)	(0.0470)	(0.0502)
MATS-2013 $\times$ distance (100 mi)			-0.108	-0.127**	-0.0769
			(0.0699)	(0.0583)	(0.0610)
MATS-2014 $\times$ distance (100 mi)			-0.152**	-0.172***	-0.121*
			(0.0748)	(0.0636)	(0.0662)
MATS-2015 $\times$ distance (100 mi)			-0.175**	-0.195**	-0.142*
			(0.0875)	(0.0786)	(0.0809)
MATS-2016 $\times$ distance (100 mi)			-0.226*	-0.246**	-0.192*
			(0.120)	(0.114)	(0.116)
Year FEs	Yes	Yes	Yes	Yes	Yes
Year X distance					Yes
Plant FEs		Yes		Yes	Yes
N	33720	33720	33720	33720	33720

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Note: Estimates are derived from OLS. Reported standard errors are clustered at the plant-level. Observations are at the unit-plant-year-level. Columns 1 and 2 present the difference-in-differences specification of the effect of MATS on propensity of a unit being converted. Estimates *do not change* following inclusion of plant fixed effects. Columns 3-4 present the triple-differences estimates of Equation 4.1, estimating the additional impact of distance under MATS.

Table A.13: Effect of Pipelines on Conversions Under MATS (WLS)

Outcome: Converted	(1)	(2)	(3)	(4)	(5)
MATS-2011	0.0172*	0.0172*	0.0299**	0.0270**	0.0233**
	(0.00887)	(0.00887)	(0.0129)	(0.0115)	(0.0118)
MATS-2012	0.0301***	0.0301***	0.0477***	0.0448***	0.0414***
	(0.0112)	(0.0112)	(0.0157)	(0.0146)	(0.0149)
MATS-2013	0.0387***	0.0387***	0.0579***	0.0550***	0.0510***
	(0.0125)	(0.0125)	(0.0173)	(0.0164)	(0.0167)
MATS-2014	0.0473***	0.0473***	0.0713***	0.0684***	0.0646***
	(0.0139)	(0.0139)	(0.0190)	(0.0182)	(0.0184)
MATS-2015	0.0783***	0.0783***	0.107***	0.104***	0.100***
	(0.0173)	(0.0173)	(0.0224)	(0.0218)	(0.0221)
MATS-2016	0.106***	0.106***	0.142***	0.139***	0.134***
	(0.0196)	(0.0196)	(0.0249)	(0.0244)	(0.0247)
MATS-2011 × distance (100 mi)			-0.152**	-0.118***	-0.0804**
			(0.0629)	(0.0384)	(0.0401)
MATS-2012 × distance (100 mi)			-0.209***	-0.175***	-0.141**
			(0.0735)	(0.0534)	(0.0557)
MATS-2013 × distance (100 mi)			-0.229***	-0.195***	-0.155**
			(0.0801)	(0.0625)	(0.0647)
MATS-2014 × distance (100 mi)			-0.286***	-0.252***	-0.214***
			(0.0871)	(0.0705)	(0.0725)
MATS-2015 × distance (100 mi)			-0.338***	-0.304***	-0.267***
			(0.0971)	(0.0835)	(0.0865)
MATS-2016 × distance (100 mi)			-0.431***	-0.397***	-0.348***
			(0.112)	(0.100)	(0.103)
Year FEs	Yes	Yes	Yes	Yes	Yes
Year X distance					Yes
Plant FEs		Yes		Yes	Yes
N	33720	33720	33720	33720	33720

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Note: Estimates are derived from WLS, where observations are weighted by the inverse number of units in a plant's fleet. Reported standard errors are clustered at the plant-level. Observations are at the unit-plant-year-level. Columns 1 and 2 present the difference-in-differences specification of the effect of MATS on propensity of a unit being converted. Estimates *do not change* following inclusion of plant fixed effects. Columns 3-4 present the triple-differences estimates of Equation 4.1, estimating the additional impact of distance under MATS.

Table A.14: Maximum Likelihood Estimates

Action	(1) MLE	(2) Weighted-MLE
No Conversion		
$\mathbb{1}(capacity \geq 25)$	-0.1658	-0.1486
	(0.0129)	(0.0128)
$\mathbb{1}(capacity \geq 25) \times \mathbb{1}(t \geq 2015)$	0.1349	0.1188
	(0.0240)	(0.0274)
Convert		
κ	7.4043	7.7163
	(1.3158)	(2.3786)
Year-Specific Fixed Costs	Yes	Yes
Log-Likelihood	-830.61	-952.88
Exp-Mean-Log-Likelihood	0.928	0.918
N	11097	11097

Note: These are the parameter estimates of the maximum likelihood estimates of Equation 4.16. Column 1 presents estimates from unweighted MLE, and Column 2 presents estimates from weighted MLE, where I weight the likelihoods by the inverse number of units in a plant's fleet. Asymptotic standard errors are derived numerically using the delta method.

Table A.15: Effect of Converted Capacity on Emissions

Pollutants	(1)	(2)	(3)
Panel A: CO_2 Equivalents (tons/MW Capacity) (mean: 1624.48)			
Plant Converted (fraction of capacity)	-1039.1***	-975.6***	-963.5***
	(218.4)	(222.9)	(207.2)
Net Generation (Kwh)			0.549***
			(0.0570)
Plant Heart Rate (mil Btu/Kwh)			4.090***
			(0.256)
N	6297	6297	6297
Panel B: NO_X (tons/MW Capacity) (mean: 2.87)			
Plant Converted (fraction of capacity)	-3.360***	-3.426***	-3.385***
	(0.599)	(0.690)	(0.683)
Net Generation (Kwh)			0.000651***
			(0.0000830)
Plant Heart Rate (mil Btu/Kwh)			0.0144***
			(0.000892)
N	6288	6288	6288
Panel C: SO_2 (tons/MW Capacity) (mean: 6.93)			
Plant Converted (fraction of capacity)	-9.653***	-8.786***	-8.739***
	(2.224)	(2.176)	(2.191)
Net Generation (Kwh)			0.00303***
			(0.000394)
Plant Heart Rate (mil Btu/Kwh)			0.00991***
			(0.00156)
N	6287	6287	6287
Year FEs	Yes		
Plant FEs	Yes	Yes	Yes
State X Year FEs		Yes	Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

Note: These are the coefficient estimates of Equation 4.17, where each panel is a separate pollutant's emissions level. Standard errors are clustered at the plant-level. Observations are aggregated to plant-year-level.

Table A.16: MATS Model Predictions

Year	ΔConverted	ΔPS ($)	ΔSCC ($)	ΔSO_2 ($)	ΔNO_X ($)	**ΔExternal Benefits**
2006	0.000	0	0	0	0	0
2007	0.000	0	0	0	0	0
2008	0.000	0	0	0	0	0
2009	0.000	0	0	0	0	0
2010	0.000	0	0	0	0	0
2011	0.009	-15,110,857,141	126,607,899	23,591,253	14,139,577	164,338,729
2012	0.019	-16,181,039,629	267,888,674	49,917,558	29,951,403	347,757,634
2013	0.030	-17,338,371,262	426,166,734	79,410,743	47,703,280	553,280,757
2014	0.042	-18,593,800,053	604,309,770	112,603,252	67,726,113	784,639,135
2015	0.063	-26,890,491,152	841,975,130	156,563,703	94,841,043	1,093,379,876
2016	0.083	-26,815,872,622	1,069,677,724	198,661,963	120,843,196	1,389,182,883
Present Value	-	-26,585,198,449	63,636,277,054	11,721,287,330	7,283,176,990	82,640,741,374

Note: Above presents the in-sample simulation results for MATS. Columns present the effect of the policy on percent of plants converted from coal to natural gas (ΔConverted), the effect on total producer surplus (ΔPS), the effect on total social cost of carbon (ΔSCC) (2015 USD), the effect on total SO_2, in 2015 dollars (ΔSO_2), the effect on total NO_X, in 2015 dollars (ΔNO_X), and the sum of these environmental externalities (Δ External Benefits = ΔSCC + ΔSO_2 + ΔNO_X). ΔSO_2 and ΔNO_X include damages from secondary pollutants produced by these emissions (such as fine particulate matter). Year-specific fixed costs are integrated out of each estimate, so that estimates are unconditional on year-specific costs. All producer surplus calculations are in present value of future expected payoffs. Producer surplus assume a price-per-mile of $1 million, so are proportional to pipeline costs in millions (i.e., for $2 million price-per-mile, multiply estimates by 2). Present value estimates simulate from post-MATS 2011 for approximately 540 periods into the future (when discounting is sufficiently close to zero), with the exception to producer surplus calculations, which just lists the present value in 2016.

Table A.17: Policy Counterfactual: MATS, with Subsidied Infrastructure

Year	ΔConverted	ΔPS ($)	ΔSCC ($)	ΔSO_2 ($)	ΔNO_X ($)	**ΔExternal Benefits**	Subsidies ($)
2006	0.003	764,861,915	28,891,864	4,963,754	3,401,786	37,257,404	32,860,963
2007	0.005	753,507,490	56,776,943	9,746,191	6,684,088	73,207,221	32,352,093
2008	0.008	742,790,296	83,701,428	14,355,893	9,852,295	107,909,617	31,854,327
2009	0.010	732,623,655	109,708,957	18,800,957	12,911,521	141,421,435	31,367,379
2010	0.013	722,942,022	134,840,727	23,089,012	15,866,606	173,796,345	30,890,971
2011	0.029	-14,868,125,764	346,285,428	61,221,893	40,016,161	447,523,482	86,334,747
2012	0.046	-15,922,654,478	572,701,341	102,076,554	65,865,201	740,643,096	90,581,575
2013	0.064	-17,062,378,892	816,452,324	146,082,459	93,682,803	1,056,217,585	95,535,153
2014	0.084	-18,297,867,550	1,080,368,257	193,753,712	123,791,069	1,397,913,038	101,334,963
2015	0.111	-24,351,849,855	1,370,127,765	246,108,585	156,884,145	1,773,120,495	123,175,132
2016	0.137	-24,359,546,783	1,643,324,878	295,458,398	188,086,278	2,126,869,554	117,166,370
Present Value	-	-24,359,546,783	78,415,247,049	14,095,526,614	8,972,775,560	101,483,549,222	1,720,374,784

Note: Above presents the in-sample simulation results for MATS. Columns present the effect of the policy on percent of plants converted from coal to natural gas (ΔConverted), the effect on total producer surplus (ΔPS), the effect on total social cost of carbon (ΔSCC) (2015 USD), the effect on total SO_2, in 2015 dollars (ΔSO_2), the effect on total NO_X, in 2015 dollars (ΔNO_X), the sum of these environmental externalities (Δ External Benefits = ΔSCC + ΔSO_2 + ΔNO_X), and the cost of the pipeline subsidies (ΔSubsidies). ΔSO_2 and ΔNO_X include damages from secondary pollutants produced by these emissions (such as fine particulate matter). Year-specific fixed costs are integrated out of each estimate, so that estimates are unconditional on year-specific costs. All producer surplus calculations are in present value of future expected payoffs. Producer surplus and pipeline subsidy amounts assume a price-per-mile of $1 million, so are proportional to pipeline costs in millions (i.e., for $2 million price-per-mile, multiply estimates by 2). Present value estimates simulate from post-MATS 2011 for approximately 540 periods into the future (when discounting is sufficiently close to zero), with the exception to producer surplus calculations, which just lists the present value in 2016.

Table A.18: Policy Counterfactual: Subsidized Infrastructure, without MATS

Year	ΔConverted	ΔPS ($)	ΔSCC ($)	ΔSO_2 ($)	ΔNO_X ($)	**ΔExternal Benefits**	Subsidies ($)
2006	0.003	764,861,915	28,891,864	4,963,754	3,401,786	37,257,404	32,860,963
2007	0.005	753,507,490	56,776,943	9,746,191	6,684,088	73,207,221	32,352,093
2008	0.008	742,790,296	83,701,428	14,355,893	9,852,295	107,909,617	31,854,327
2009	0.010	732,623,655	109,708,957	18,800,957	12,911,521	141,421,435	31,367,379
2010	0.013	722,942,022	134,840,727	23,089,012	15,866,606	173,796,345	30,890,971
2011	0.015	714,789,423	159,171,356	27,233,802	18,726,495	205,131,653	30,424,831
2012	0.017	706,955,029	182,733,771	31,241,406	21,495,051	235,470,228	29,968,695
2013	0.019	699,412,073	205,559,176	35,117,575	24,175,945	264,852,696	29,522,308
2014	0.021	692,138,167	227,677,153	38,867,754	26,772,663	293,317,570	29,085,420
2015	0.024	684,147,023	249,085,602	42,491,597	29,284,846	320,862,045	28,657,790
2016	0.026	676,450,095	269,814,211	45,994,564	31,716,025	347,524,800	28,239,181
Present Value	-	676,450,095	7,861,866,110	1,332,002,894	923,179,423	10,117,048,427	602,375,132
Marginal Value Per-Mile	-	-	2,051,973	347,657	240,953	2,640,582	157,222

Note: Above presents the in-sample simulation results for MATS. Columns present the effect of the policy on percent of plants converted from coal to natural gas (ΔConverted), the effect on total producer surplus (ΔPS), the effect on total social cost of carbon (ΔSCC) (2015 USD), the effect on total SO_2, in 2015 dollars (ΔSO_2), the effect on total NO_X, in 2015 dollars (ΔNO_X), the sum of these environmental externalities (Δ External Benefits = ΔSCC + ΔSO_2 + ΔNO_X), and the cost of the pipeline subsidies (ΔSubsidies). ΔSO_2 and ΔNO_X include damages from secondary pollutants produced by these emissions (such as fine particulate matter). Year-specific fixed costs are integrated out of each estimate, so that estimates are unconditional on year-specific costs. All producer surplus calculations are in present value of future expected payoffs. Producer surplus and pipeline subsidy amounts assume a price-per-mile of $1 million, so are proportional to pipeline costs in millions (i.e., for $2 million price-per-mile, multiply estimates by 2). Present value estimates simulate from 2006 (the start of the counterfactual subsidies) for approximately 540 periods into the future (when discounting is sufficiently close to zero), with the exception to producer surplus calculations, which just lists the present value in 2016. Estimates for the marginal value per-mile of pipeline are calculated for relevant outcomes based on the approximately 3,831 of mileage necessary to connect each coal plant (i.e., marginal value = present value/3,831).

APPENDIX B

SUPPLEMENTAL CONTENT

B.1 Fuzzy Matching Accidents to Owners

Due to an absence of VIN or household-level information in the FARS dataset, I make use of additional information in the TTI data to "fuzzy" match fatal vehicle accident victims to households in the DMV dataset. This process involves two layers of "fuzzy" matching, in which I first match key TTI information to FARS observations, and then match fatal accident victims to vehicle owners in the DMV data.

The Texas fatal vehicle accidents in the FARS dataset are a small subset of the TTI dataset, which records information on all vehicle accidents in Texas, whereas FARS records only accidents involving at least one fatality. I can match these two datasets together using information on the time and location of the accident, as well as the make, model, and year of the cars involved. I merge these two datasets together in order to supplement the FARS dataset with the vehicle owner names recorded in the TTI data. With this supplemented information, I will be able to more accurately match victims of fatal accidents to households in the DMV.

Due to minor inconsistencies in the data, I must *fuzzy* match the TTI to FARS. For instance, geo-coordinates, as well as the exact time of the crash, may differ slightly between datasets. My first condition for a vehicle listed in FARS to be a candidate match for a vehicle listed in TTI is an exact match on vehicle model year. Second, I require the crash coordinates to differ by no more than 4,000 meters in each dataset. With this much smaller set of candidate TTI vehicles, I sort by the difference in crash times reported. After this procedure, if I am left with a non-unique match on time, I also sort by physical distance of crash. It is rarely the case that I am left with a non-unique match following these conditions. When the result here is non-unique, it is more than likely the case that these vehicles were involved in the same accident and each vehicle had the same model year. To narrow it down to a unique match, I then choose the match with the longest common substring in terms of vehicle make.[1]

I use this finalized set of fatal car accidents to fuzzy match to DMV records. I require an exact match on the (non-unique) 12-digit substring of the 17-digit VIN number, and then fuzzy match on first and last name jointly, using the Levenhstein distance, with a critical value of 0.25[2]. I also add a condition on dates, requiring candidate vehicle matches in the DMV dataset to be owned prior to the crash. Often, I am still left with a list of candidate matches in the dataset. Finally, I narrow my results down to the household who's address is closest to the accident. This last criteria is based on the common statistic that most vehicle accidents occur close to home.[3]

B.2 Factors Contributing to a Trend Break in the Control Group

One might be concerned with the trend break for the control group at around the time of the accident in Figure A.2. The median crash time in my sample occurred around the summer of 2008, so it is possible that this drop in vehicle size was due to the high gas prices. We would then suspect that the treatment group would dip downwards as well, if it were not for the neighbor dying.

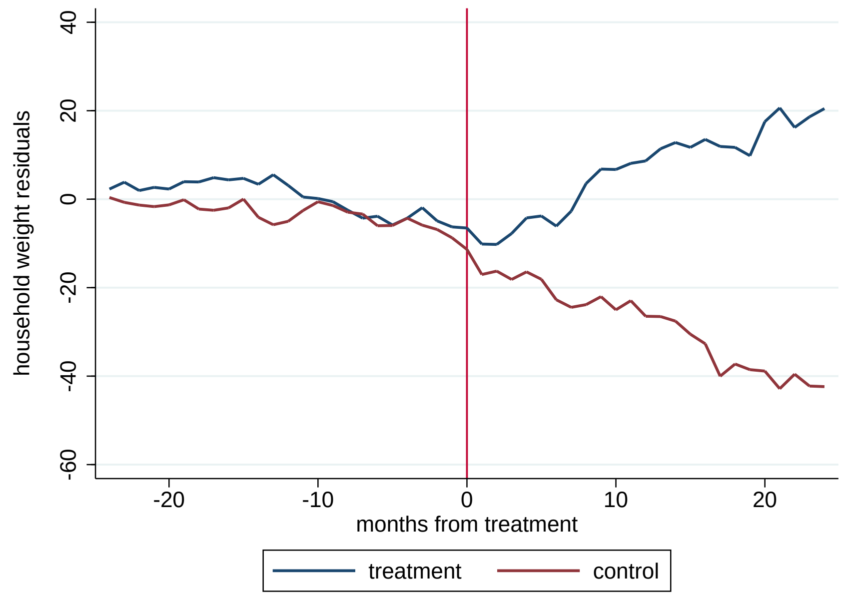
Next, I demean household vehicle weight by month-of-sample, and plot the residuals in the same manner. The results are in Figure B.1. Accounting for any systematic changes over time eliminates the trend break in the control group, suggesting, strictly time-specific factors affecting the control group. This implies, assuming my identification strategy is valid, that these factors would have affected the treatment group in the same manner, absent treatment.

[1] Vehicle make in the two datasets may differ slightly. For instance "Ford" versus "Ford Motor Company".

[2] The Levenhstein distance is the minimum number of insertions, deletions, or substitutes of characters required to change one string into another. For a joint twelve letter first and last name, a Levehnstein distance of 3 requires a critical value of 0.25.

[3] For example, a 2001 Progressive Insurance Survey found that 52 percent of car crashes occur within five miles or less from home, and 77 percent occur within 15 miles from home.

Figure B.1: Effect of Neighbor's Fatality on Household Vehicle Size, Conditional on Month-of-Sample Fixed Effects



I also examined whether the trend break was being driven by gasoline prices or employment, but I was unable to find any one variable that fully accounted for this trend break. As it seems to be a time-specific factor, it will not bias my estimates, as my identification strategy is robust to such factors.

B.3 Additional Outcomes

I test whether a neighbor dying in a vehicle accident has an effect on other attributes of consumer decision making beyond increasing the size of their vehicles. Figure B.2 presents the analogue to Figure A.2 for three additional outcomes.[4]

In Figure B.2, panel a, I examine the extent to which this transition to heavier vehicles is translated

[4]Additionally, I look at whether the timing of purchase changed and the value of the vehicle stock—which I define as the average price that VIN10 was sold for in that month. The timing only changes slightly, and insignificantly (for the two neighbor sample). The vehicle value measure was not very precise and lead to noisy estimates. This is mostly due to insufficient number of the same VIN10s being sold in any given month

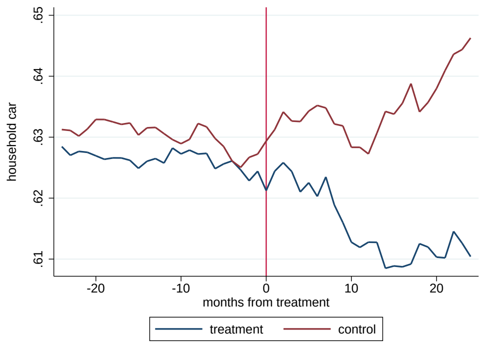
into a switch away from passenger cars. This figure illustrates that treated households decrease their consumption of passenger cars. This decrease in probability of owning a passenger car implies an increase in adoption of light trucks.

In panel b, I look at the effects on vehicle safety, as measured by the fraction of vehicles a household owns that have at least one airbag. This variable was constructed using a string variable in my data which describes the restraint system in a vehicle. This description will normally include information regarding the types of seatbelt and airbag systems installed in the car. I assign a one to vehicles that contain “airbag”, “Airbag”, or “AIRBAG” in this description, and a zero otherwise. Obviously this variable will have a high baseline outcome, as most cars in the period examined have airbags, however, in the figure, following treatment, it appears that the treatment group increases their relative share of vehicles with at least one airbag. Though the crossed paths in the pre-period could be a reason for concern.

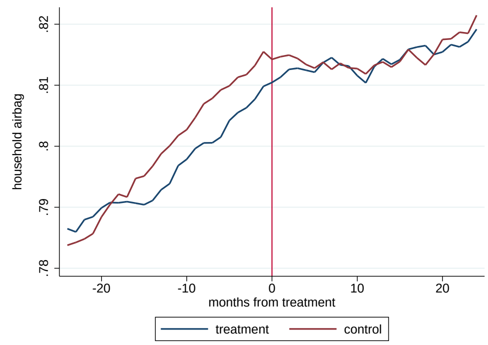
Panel c illustrates the effect on household average combined MPG. Because of the relationship between vehicle size and fuel efficiency, we would expect to see an effect here. The figure illustrates that the control and treatment group seem to track each other well in the pre-period. Following treatment, the treated group shows a significant drop in MPG.

Figure B.2: Effect of Neighbor's Fatality on Other Outcome

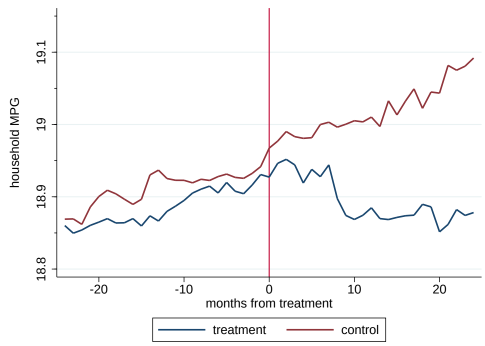
(a) Household Owns a Passenger Car



(b) Household Fraction of Fleet with Airbag



(c) Household MPG



In Table B.1, I present the estimated treatment effects for these additional outcomes. As in Table A.2, these effects are estimated using the nearest two neighbors on the same street. The estimates in panel a are the results of a linear probability model on whether the household owns at least one passenger car. These estimates imply a reduction in probability of owning a passenger car of about two percentage points. The estimated effect on household airbags implies that a neighbor's fatality in an accident produces a near one percentage point increase in proportion of vehicles with airbags. Such a small effect should not be surprising since most households already own vehicles with airbags. The effect of a neighbor's fatality on average vehicle fleet MPG is about a 0.1 reduction. Since this is the effect on household average MPG, this implies about a 0.2 reduction in MPG for a two-vehicle household replacing only one vehicle in the post period.

Table B.1: Effect of Neighbor's Fatality in a Vehicle Accident

	(1)	(2)	(3)	(4)	(5)
Panel a: Passenger Car					
neighbor died	-0.0230***	-0.0210**	-0.0212**	-0.0173***	-0.0171***
	(.0087)	(0.0086)	(0.0086)	(0.0062)	(0.0062)
N	1,003,162	1,003,162	1,003,162	1,003,162	1,003,162
Panel b: Airbag					
neighbor died	0.0029	0.0052	0.0055	0.0072*	0.00702*
	(0.00579)	(0.00581)	(0.00580)	(0.00416)	(0.00411)
N	1,003,162	1,003,162	1,003,162	1,003,162	1,003,162
Panel c: MPG					
neighbor died	-0.131*	-0.111	-0.111	-0.0895**	-0.0744*
	(.0706)	(0.0692)	(0.0692)	(0.0438)	(0.0437)
N	972,893	972,893	972,893	972,893	972,893
Controls		Yes	Yes		
Month $\times$ Year FE			Yes	Yes	Yes
Household FE				Yes	Yes
County Time Trend					Yes

Standard errors in parentheses. Standard errors are clustered at the street level.

* $p < 0.1$, ** $p < .05$, *** $p < .01$

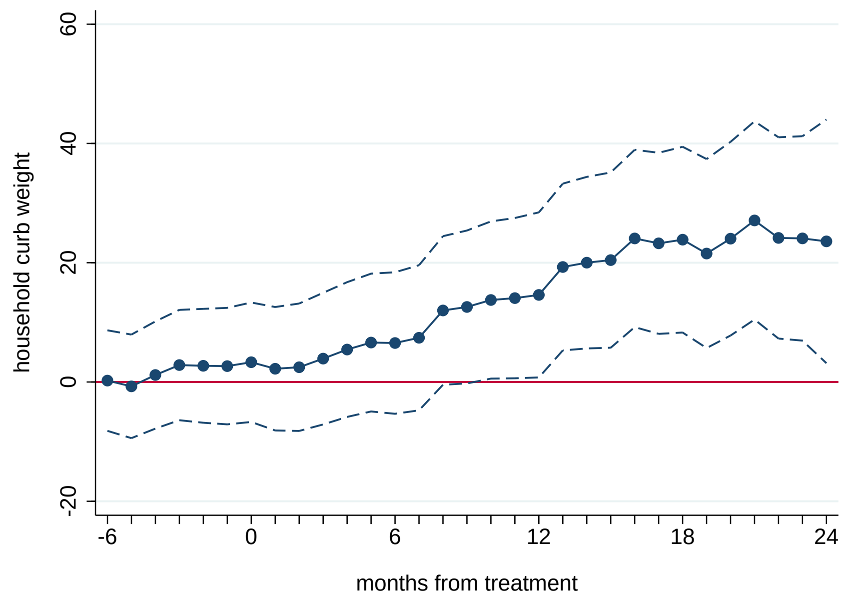
B.4 An Alternative Control Group

In this section, I look at an alternative control group to check if my definition of the control group as the *survivors* makes a difference. The results from Table A.3 form the basis of this control group. As neighbors further away from the involved household are added to my sample, the treatment effect becomes attenuated. This implies that neighbors further away (but on the same street) are less effected by treatment.

I construct an alternative control group by assigning the tenth or greater treated neighbors to the control, while the treatment group remains the nearest two neighbors to the accident involved household. It is important to note that these *far away* neighbors are still treated. Therefore, I may run into the same problem as my original control group—both increase weight following their neighbor's accident.

The analogue to Figure A.3 for the new control group is presented in Figure B.3 below. Comparing these estimates to those with the original control group demonstrates a (very) near-identical relationship. Similar to the original control group estimates, these estimates imply an overall treatment effect of between 20–25 pounds. The zero effect in the pre-periods also demonstrates that this control group tends to track the treatment group well.

Figure B.3: Dynamic Estimates for Weight: Treatment Neighbors $\geq$ 10 as Control



B.5 Alternative Mechanisms

In this paper, I argue that consumers respond to a salient vehicle accident of a neighbor by increasing their relative vehicle size due to an increased preference for safety. Such a response is consistent with a vehicle arms race. However, it is possible that consumers are increasing their vehicle size due to reasons other than preferences for safety. One alternative hypothesis is that consumers observe a neighbor's death, which triggers a belief that life is short, and thus they should treat themselves to a larger, more luxurious vehicle.

In this section, I test whether consumers are responding for reasons other than preferences for safety, by examining the effect of neighbor's fatality on vehicle price, conditional on weight. The idea is that, other mechanisms would more than likely lead consumers to not only respond in terms of vehicle size, but also, additional vehicle bells and whistles, which should be correlated with the vehicle's price. For example, the "life is short" hypothesis suggests that I observe a neighbor dying, and decide to by something more luxurious, which might be correlated with vehicle size. These additional features should be reflected in vehicle prices.

To test for these other possible responses, I will look at the transaction prices of the vehicles in a household's fleet, and test whether these households increase the value of their fleet by means other than purely increasing the size of the fleet. That is, post-neighbor fatality, do households increase the value of their fleet, conditional on fleet size?

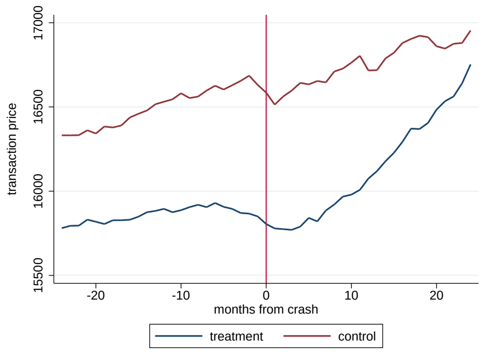
I've constructed two new variables, that vary by household at a monthly frequency. *Transaction price* is the household's average price paid for each vehicle in the fleet. This outcome only changes following a purchase, and therefore, remains constant across most observations. *Car value* calculates the average value of the household's fleet, assuming a 1.3% depreciation rate per month.[5]

Figure B.4 plots the value of a vehicle fleet over time, according to these two separate metrics. Both measures show an increase in the treatment group's fleet value following the accident. This makes since due to the positive relationship between a vehicle's size and its price.

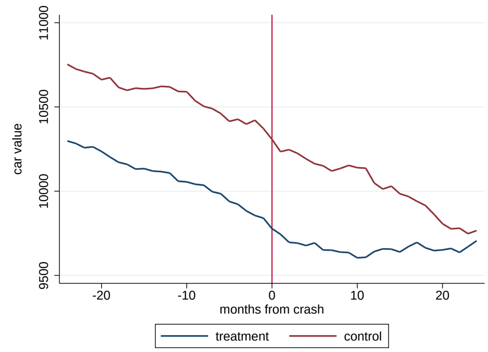
[5]This corresponds to the estimated average depreciation rate of a vehicle, as reported by *Edmunds.com*.

Figure B.4: Effect of Neighbor's Fatality on Fleet Price

(a) Household Transaction Price



(b) Household Car Value



In Table B.2, I present the results of the difference-in-differences regression, with fleet value as the outcome variable. I show these estimates before and after controlling for weight. Though not statistically significant, the estimates suggest a positive effect of a neighbor's death on vehicle value. These estimates decrease in size after controlling for fleet weight.

Table B.2: Effect of Neighbor Dying on Vehicle Price

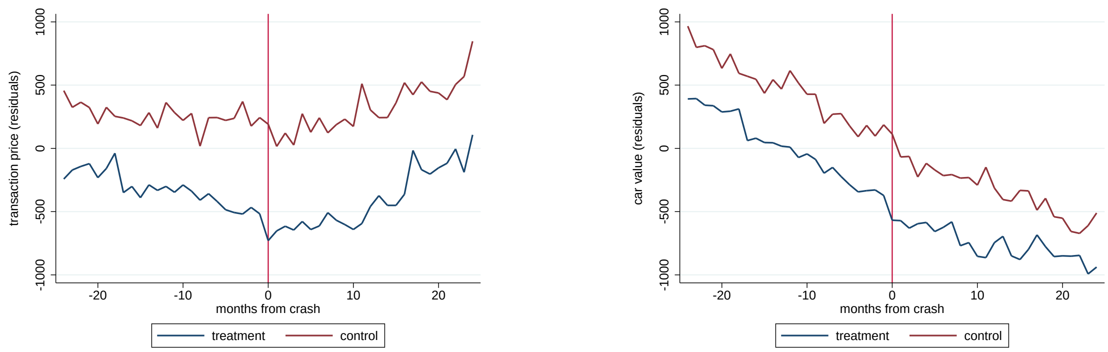
	(1)	(2)	(3)	(4)
Panel b: Transaction Price				
neighbor died	135.4	-167.6	64.59	-63.28
	(206.9)	(209.4)	(94.07)	(89.59)
weight		6.171***		4.584***
		(0.0916)		(0.118)
Panel b: Car Value				
neighbor died	159.3	-16.49	99.76	-2.461
	(149.7)	(154.7)	(97.63)	(103.8)
weight		4.525***		4.352***
		(0.0679)		(0.109)
Yr-Month FE			Yes	Yes
HH FE			Yes	Yes
N	884,416	540,628	884,416	540,628

Standard errors in parentheses. Standard errors are clustered at the street level.
* $p < 0.1$, ** $p < .05$, *** $p < .01$

Next, I present these results graphically. I regress the fleet value variables on household average weight. I plot the residuals from this regression in Figure B.5. This Figure is the analog to Figure B.4, after controlling for fleet weight. This figure suggests that the increase in fleet value seen in Figure B.4 was primarily due to an increase in size.

Figure B.5: Effect of Neighbor's Fatality on Fleet Price, Conditional on Fleet Size

(a) Household Transaction Price Residuals (b) Household Car Value Residuals



If households respond to a neighbor's death by purchasing additional features of a car, other than weight, we would suspect to see an increase in vehicle value by means other than that of just vehicle size. Therefore, this evidence suggests that other possible mechanisms, such as this "life is short" example, are not a driving force, further supporting the notion that consumers increase their preferences for safety—behavior consistent with a vehicle arms race.

B.6 Optimal Tax Policy

The results in Section 2.7 produce important implications for Pigovian tax policy. Of particular interest is the result illustrated in Figure A.8 that suggests an arms race can be reversed with a sufficiently high gas price. In this section, I directly estimate the optimal gas tax such that the arms race effect is fully internalized. Specifically, I derive the period-specific gas tax that will incentivize the arms race group to consumer in the same manner as the no arms race group.

Figures B.6 and B.7 plot the optimal tax rates that force the arms race group to consume identical vehicle sizes as the no arms race group, when the no arms race group faces a given (tax-free) price of gasoline. Of particular interest is the result in Figure B.6, Panel b. The results in Section 2.7 imply that policy may be able to exploit arms race preferences, and these results confirm this. The results indicate, to fully internalize an arms race in the light county under low gas prices, tax rates should initially be set high, and later reduced. The initial high gas price kick starts the reverse arms race. Once consumers become less concerned with the shrinking fleet size, tax rates may be reduced.

Figure B.6: Optimal Gas Tax to Internalize the Arms Race (fuel price = $1.665)

(a) Heavy County (b) Light County

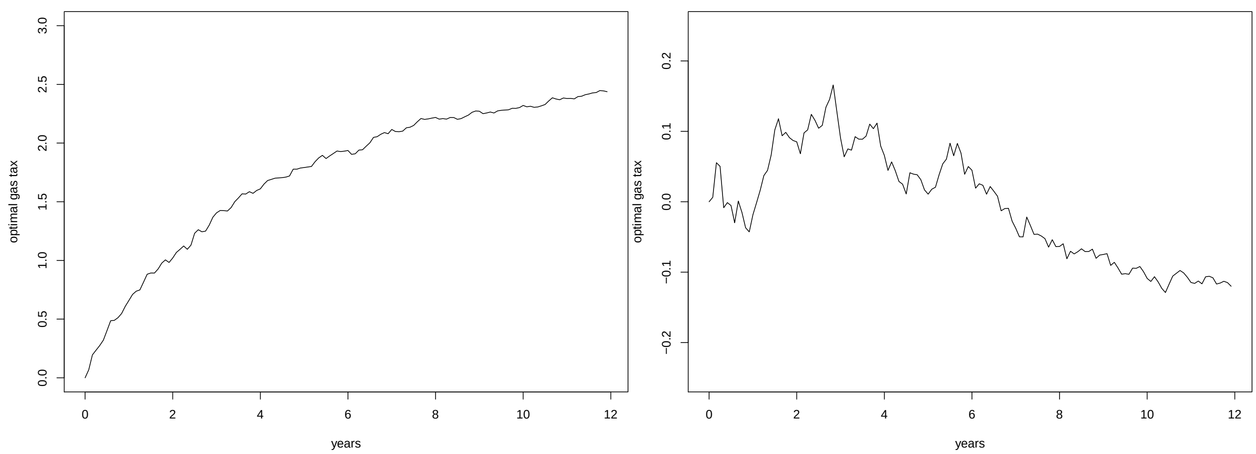
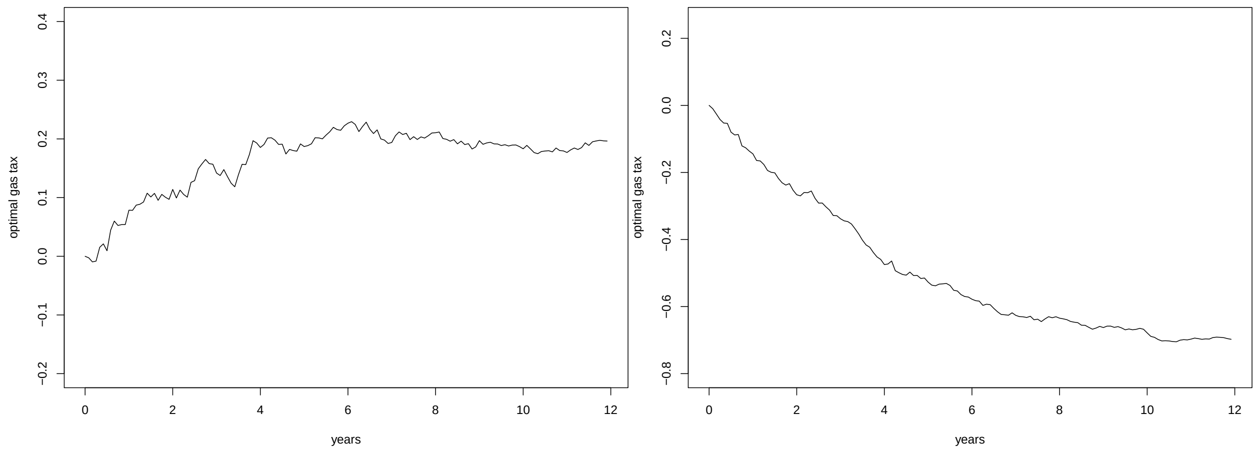


Figure B.7: Optimal Gas Tax to Internalize the Arms Race (fuel price = $4)

(a) Heavy County (b) Light County



B.7 Counterfactuals with Equilibrium Prices

In this section, I run the same counter-factuals, but now allowing firms to re-optimize prices in every period. Of course, in the vehicle market, manufacturers are multi-product firms. For simplicity, I treat each unique *make* as its own company, and optimize prices over models.

I follow the standard model of multi-product firm optimization (e.g., Bresnahan, 1987). Suppose there are $\mathcal{F}$ firms (makes, in my case), and each firm, f, maximizes profits over a set of own-products, J_{f}. The firm is assumed to maximize profits according to:

$$\max_{\{p_j\}_{j\in J_{\text{f}}}} \left\{ \Pi_{\text{f}} = \sum_{j\in J_{\text{f}}} s_j \cdot (p_j - c_j) \right\} \tag{B.1}$$

where s_j is the choice probability for car j (a function of price), p_j is the price, and c_j is constant marginal cost of car j. Profit maximization then implies the following first order condition for car j belonging to firm f:

$$s_j + \sum_{k \in J_{\mathrm{f}}} \frac{\partial s_k}{\partial p_j} \cdot (p_k - c_k) = 0$$

This becomes a set of J equations, where J is the total number of vehicles in the market. Following the framework of Nevo (2000), I can write this in matrix form as:

$$S - \Omega \cdot (P - C) = 0$$

where P is the vector of J prices, C the vector of marginal costs, S is the vector of choice probabilities, and Ω is a $J \times J$ matrix, defined as follows:

$$\Omega_{jk} = \begin{cases} -\frac{\partial s_j}{\partial p_k}, & \text{if } \exists \text{f } st.\{j,k\} \subseteq J_{\mathrm{f}} \\ 0, & \text{otherwise} \end{cases}$$

Under the assumption that the prices observed in the data are the Nash equilibrium outcome of Bertrand competition, I can calculate the implied vector of marginal costs as:

$$C = P - \Omega^{-1} \cdot S \tag{B.2}$$

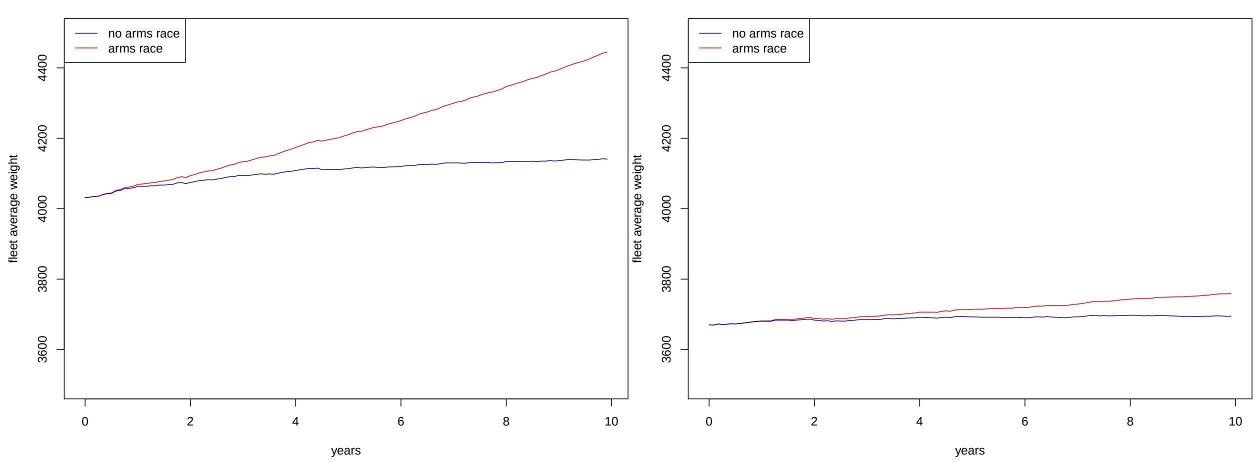
As I am using average transaction price by make-model-year, in a given county, I will compute this marginal cost vector in each county, given the choice set for model-year 2009 vehicles. I simulate purchases in the same manner as in the previous section, only now, allowing firms to optimize prices at the beginning of each period. I assume that marginal costs remain constant throughout the simulation.

The simulation proceeds as follows: beginning at $t = 0$, the observed price vector is as observed in the data, and as used in all simulated periods in the previous exercise. Marginal costs are the solution to Equation B.2, and randomly drawn consumers make their purchase decisions accordingly. At $t = 1$, firms observe higher (lower) mean county weights and respond to the increased (decreased) willingness to pay for pounds. The firms enter Bertrand competion according to Equation B.1. Bertrand-Nash equilibrium prices are calculated, and these are the prices that consumers observe in that period. Simulated consumers observe new equilibrium prices, new mean weights, and fuel prices, and make their decision accordingly.

Figure B.8 is the analog to Figure A.7, when consumers face new equilibrium prices in each period. The results are qualitatively the same. The weight differentials are significantly smaller than the partial equilibrium outcomes for both counties. This makes since, as rational firms should increase mark-ups for large vehicles as consumers increase their willingness to pay for them.

Figure B.8: Counterfactual Mean County Weights with Equilibrium Prices (fuel price = $1.665)

(a) Heavy County ($\Delta\bar{w}_T = 304$) (b) Light County ($\Delta\bar{w}_T = 65$)



With low gas prices at equilibrium vehicle prices, consumers in the heavy county still increase their vehicle size by a meaningful 304 pounds over the course of 10 years, relative to no arms race. This effect amounts to about 7.6% over the initial mean weight. On the other hand, the light county very slowly increases the gap between arms race and no arms race mean weights. After 10 years, the difference between the two groups in the light county is only 65 pounds. This amounts to only 1.7% over the initial mean weights, following a decade of purchases.

Figure B.9 shows the effect under high gas prices. Again, these results are qualitatively similar, though smaller in scale, to those in the partial equilibrium analysis. Under sustained high gas prices for 10 years, the heavy county has an effect of about 122 pounds, or 3% of initial mean weights.

Figure B.9: Counterfactual Mean County Weights with Equilibrium Prices (fuel price = $4)

(a) Heavy County ($\Delta\bar{w}_T = 122$) (b) Light County ($\Delta\bar{w}_T =$ -84)

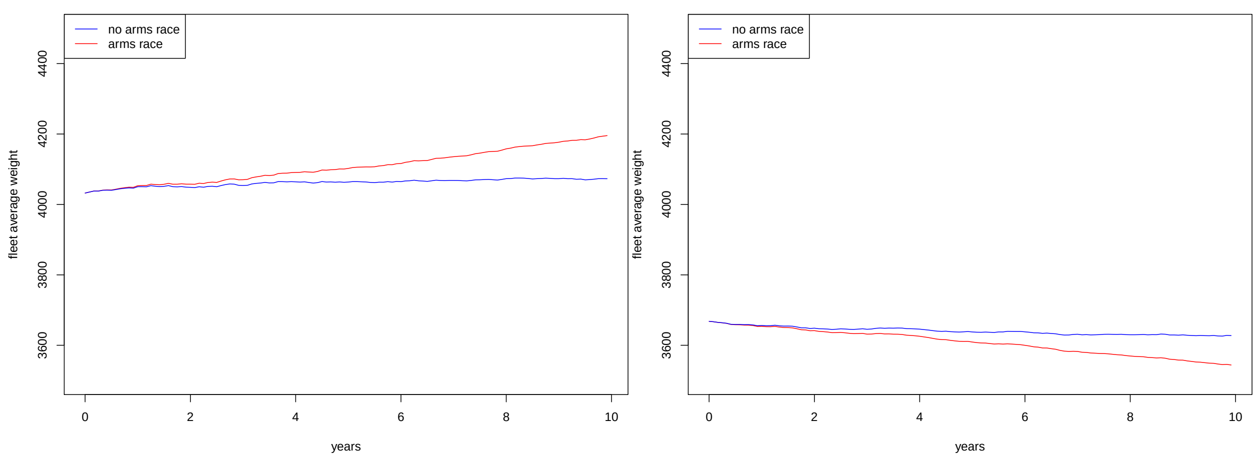


Figure B.10: Log-Difference in CO_2/mile Emissions of New Purchases (fuel price = \$1.665)

(a) Heavy County ($\%\Delta \text{CO2}_T = 10\%$) (b) Light County ($\%\Delta \text{CO2}_T = 2.5\%$)

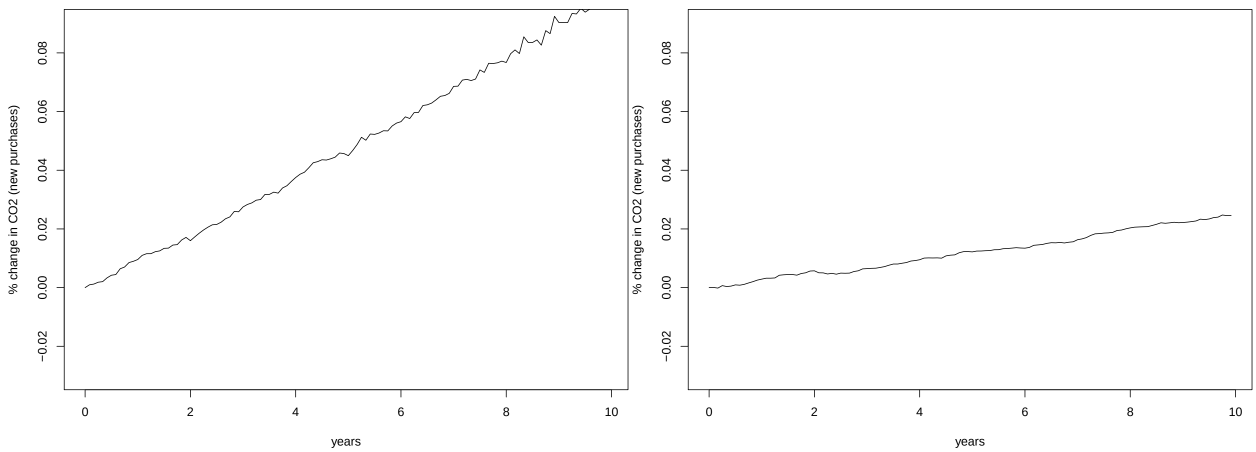
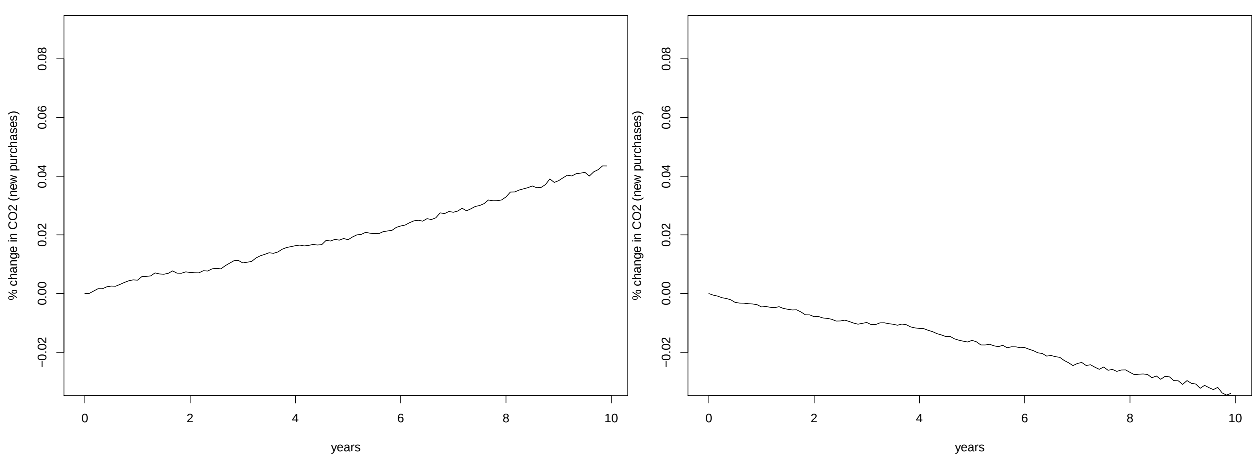


Figure B.11: Log-Difference in CO_2/mile Emissions of New Purchases (fuel price = \$4)

(a) Heavy County ($\%\Delta \text{CO2}_T = 4.3\%$) (b) Light County ($\%\Delta \text{CO2}_T = \text{-}3.4\%$)



B.8 Alternative Instrumental Variable

FEMA's assignment of FHBMs in the 1970s map closely into the communities that ultimately receive a FIRM. Our data aggregate community up to the county-level, and our sample suggests that nearly 99 percent of counties with an FHBM ultimately receive a FIRM; though there is still some variation between FHBM and FIRM proportions within county. In this paper, we proxy treatment with FHBM assignment, as these counties were identified as risky by FEMA, and thus, were more likely to ultimately enroll into the NFIP. We exploit the timing of FIRM assignment and use the interaction of *FHBM* and *post-FIRM* as our instrument for NFIP enrollment. Not separately controlling for *post-FIRM* produces estimates as deviations from baseline fixed effects levels, rather than as deviations from post-FIRM. As our *postFIRM-FHBM* instrument for endogenous NFIP ($treat \times post$) we argue that this deviation from baseline fixed effects is the variation we want to isolate.[6]

In this section, we present our estimates when using *postFIRM* as our instrument for the entire sample of counties, rather than *postFIRM* for the "more likely to be treated," *FHBM*, group (i.e., the interaction $postFIRM \times FHBM$). Note that this will change not only our reduced form equation, but the first stage as well. Therefore, of primary interest is the scaled two-stage least squares estimates. Our estimates of this alternative first stage are presented in Table B.3.

Table B.3: First Stage: Effect of FIRM on NFIP Enrollment

Post-NFIP	(1)	(2)	(3)
postFIRM	0.761***	0.761***	0.761***
	(0.0134)	(0.0134)	(0.0134)
County FE	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes
Controls		Yes	Yes
Declared Disaster Controls			Yes
N	64472	64472	64472

* $p < 0.1$, ** $p < .05$, *** $p < .01$

These estimates are significantly smaller than those from our original instrument in Table A.8. This suggests that our original instrument—with a first stage estimate of 0.95—map into NFIP enrollment much closer than FIRM alone—an estimate of 0.76. Next, we estimate our reduced

[6]Note that *postFIRM-FHBM* and *postFIRM* are extremely collinear, and split our estimates when including both in our primary estimation.

form equation with *postFIRM* as our instrument. This is simply the analogue to Equation 3.2, and Table A.9 results, but substituting *postFIRM* for *postFIRM-FHBM*. The results are in Table B.4.

Table B.4: Effect of FIRM on Migration

Migration Outcome	(1)	(2)	(3)	(4)
Panel A: Log- Population				
postFIRM	0.0370***	0.0345***	0.0333**	0.0333**
	(0.0131)	(0.0124)	(0.0135)	(0.0135)
Leading Treatment			-0.00613	-0.00615
			(0.00882)	(0.00882)
Panel B: Log- Non-Migrants				
postFIRM	0.0391***	0.0362***	0.0354**	0.0354**
	(0.0134)	(0.0126)	(0.0138)	(0.0138)
Leading Treatment			-0.00433	-0.00434
			(0.00903)	(0.00903)
Panel C: Log- Migrants				
postFIRM	0.0268*	0.0267*	0.0239	0.0240
	(0.0148)	(0.0143)	(0.0156)	(0.0156)
Leading Treatment			-0.0142	-0.0142
			(0.0119)	(0.0119)
N	64472	64472	64472	64472
County FE	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes
Controls		Yes	Yes	Yes
Declared Disaster Controls				Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

These results produce a smaller reduced form estimate than our original instrument. This is to be expected with a smaller first stage. Estimates for population indicate a 3.3 to 3.7 percent effect. Scaling by the first stage of 0.76, our results indicate an effect of about 4.3 to 4.9 percent. These results are near identical to our primary estimates, and the same holds true for our other migration outcomes. Therefore, exploiting *FHBM* adds little to our overall analysis, but only produces a first stage closer to one.

B.9 Exogeneity of County Flood Risk

As our NOAA annual floods data begin in 1996, we are not able to use pre-NFIP flood risk for our analysis in Section 3.6. Using average flood risk over these periods with high enrollment into flood insurance may produce biased estimates in our triple-diff specification if NFIP systematically alters a county's flood risk. For example, if NFIP enrollment causes a community to invest in flood mitigation infrastructure, we would suspect our heterogeneous treatment effects of NFIP by flood risk to be downward bias. That is, our measure of risk would reveal the communities which invested the least in flood mitigation infrastructure, presumably contributing a lower amount to population.

In this section we provide evidence that our triple-diff estimates are consistent by examining the relationship between NFIP enrollment and flood occurrences. Note that we do not anticipate these estimates to be *causal*, as it is well-documented in the literature that major floods increase flood insurance take-up (Gallagher, 2014; Kousky and Shabman, 2014). In this scenario, our estimates of the effect of NFIP enrollment on flood risk will be downward biased (away from zero), as floods will *Granger* cause enrollment, followed by reductions in floods. In a difference-in-differences framework, this is represented by a positive effect in the pre-NFIP periods. Thus, rejecting a negative effect of NFIP on floods provides evidence that our estimates are consistent.

We estimate the following equation:

$$floods_{cst} = \delta \cdot \textit{post-NFIP}_{cst} + \lambda_{st} + \gamma_{cs} + \varepsilon_{cst} \tag{B.3}$$

where $floods_{cst}$ is the annual flood episodes, reported by NOAA, for a given county-state-year, cst. *post-NFIP*$_{cst}$ reports the fraction of a county enrolled into NFIP, and λ_{st} and γ_{cs} are state-by-year and county-state fixed effects, respectively. The estimate for equation B.3 is reported in Table B.5. Standard errors are clustered by county.

Table B.5: Effect of NFIP on County Floods

	floods
Post-NFIP	-0.150
	(0.0960)
County FE	Yes
State X Year FE	Yes
N	51712

* $p < 0.1$, ** $p < .05$, *** $p < .01$

These estimates suggest reduction of 0.15 in future floods from the NFIP, however, this relationship is not statistically significant. Given that this estimate is plausibly biased away from zero, this provides strong evidence that using historical floods from this time period should not bias our estimates. To further test the consistency of our estimates in Section 3.6, we now difference out these estimated effects of NFIP on floods, and use residuals in our triple-diff estimates. These residuals our, by construction, orthogonal to NFIP treatment. In the same manner as our primary estimates, we average the residual floods by county and use this cross-sectional representation of flood risk to estimate the differential effects of NFIP across counties. We present the two staged least squares estimates (using average county flood residuals as our instrument for average floods) in Table B.6. These estimates are marginally larger in magnitude, suggesting the correctly anticipated direction of bias, though not signficantly different from those in Table A.10.

Table B.6: Heterogeneous Effects of NFIP on Migration Using Residualized Floods (2SLS)

	Migration Outcome		
	(1)	(2)	(3)
Panel A: Log- Population			
postFIRM-FHBM	0.00707	0.000585	0.000620
	(0.0256)	(0.0247)	(0.0247)
Annual Floods × postFIRM-FHBM	0.0282*	0.0338**	0.0339**
	(0.0151)	(0.0141)	(0.0141)
Panel B: Log- Non-Migrants			
postFIRM-FHBM	0.00508	-0.00214	-0.00212
	(0.0269)	(0.0261)	(0.0261)
Annual Floods × postFIRM-FHBM	0.0317*	0.0377**	0.0377**
	(0.0164)	(0.0154)	(0.0154)
Panel C: Log- Migrants			
postFIRM-FHBM	0.00921	0.00886	0.00894
	(0.0271)	(0.0264)	(0.0264)
Annual Floods × postFIRM-FHBM	0.0131	0.0164	0.0165
	(0.0156)	(0.0149)	(0.0149)
N	67559	67559	67559
County FE	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes
Year X Floods	Yes	Yes	Yes
Controls		Yes	Yes
Declared Disaster Controls			Yes

* $p < 0.1$, ** $p < .05$, *** $p < .01$

These estimates are marginally larger in magnitude as those in Table A.10, suggesting the correctly anticipated direction of bias. However, these differences are not statistically significant. Therefore, we are confident that the use of in-sample floods to evaluate a county's average risk level should not significantly bias our estimates.

B.10 Proof of Proposition 1

Equation 4.2 defines plant i's per-period profits. From this, define plant i's per-period costs at time t as:

$$c_{it} = \sum_{j \in J_i} c_j(f_{jt}) + k_i \times \mathbb{1}(f_{jt-1} = coal\ \forall j \in J_i \wedge f_{j't} = NG \text{ for some } j' \in J_i)$$

with set of units J_i, fuel sources $\{f_{jt}\}_{j \in J_i}$, and $f_{jt} \in \{coal, NG\}$ for all $j \in J_i$ in all periods, t.

We can define the marginal cost of converting some unit, $j' \in J_i$, as the difference in plant costs when j' uses NG versus $coal$, holding other fuel sources constant. That is,

$$MC_{j't} = c_{it}|_{f_{j't}=NG} - c_{it}|_{f_{j't}=coal}$$

There are two cases which must be satisfied for fixed costs of conversion, k_i, to enter plant costs in period t. Mainly, 1) all fuel sources in the previous period were $coal$ ($f_{jt-1} = coal\ \forall j \in J_i$), and 2) at least one unit is converted to natural gas at t ($f_{j't} = NG$ for some $j' \in J_i$).

Case 1: when $f_{jt-1} = NG$ for some $j \in J_i$, the result follows trivially, as the plant does not incur k_i.

Case 2: $f_{jt-1} = coal\ \forall j \in J_i$

We can write the plant costs evaluated at $f_{j't} = NG$ and $f_{j't} = coal$, respectively as:

$$c_{it}|_{f_{j't}=NG} = \sum_{j \in J_i \setminus j'} c_j(f_{jt}) + c_{j'}(NG) + k_i$$

$$c_{it}|_{f_{j't}=coal} = \sum_{j \in J_i \setminus j'} c_j(f_{jt}) + c_{j'}(coal) + k_i \cdot \mathbb{1}(f_{jt} = NG \text{ for some } j \in J_i \setminus j')$$

This implies the following marginal costs for the plant:

$$
\begin{aligned}
MC_{j't} =& c_{j'}(NG) - c_{j'}(coal) + k_i \cdot (1 - \mathbb{1}(f_{jt} = NG \text{ for some } j \in J_i \setminus j')) \\
=& c_{j'}(NG) - c_{j'}(coal) + k_i \cdot \mathbb{1}(f_{jt} = coal \; \forall j \in J_i \setminus j')
\end{aligned}
$$

From *Case 1* and *Case 2*, it follows that, for any $j \in J_i$:

$$
\begin{aligned}
MC_{ijt} =& -\big(\mathrm{c}_j(NG) - \mathrm{c}_j(coal)\big) \\
&- \kappa_i \cdot \mathbb{1}\big(f_{j't-1} = coal \; \forall j' \in J_i \quad \wedge \quad f_{j't} = coal \; \forall j' \neq j, \; j' \in J_i\big)
\end{aligned}
$$

$\square$

B.11 Effect of Conversions on Plant Generation

In this section, I directly test whether plant annual generation levels change significantly following a conversion. Rather than including generation on the right hand side in Equation 4.17 as a control, and testing whether δ is significantly affected by its inclusion (which Table A.15 rejects), here I directly examine the relationship between conversions and annual electricity generation, by placing net generation on the left hand side. Table B.7 reports these estimates, which rejects any significant effect of a coal-to-natural gas conversion on a plant's annual electricity generation.

Table B.7: Effect of Conversions on Plant Generation Levels

Net-generation (Kwh)	(1)	(2)	(3)
Plant Converted (fraction of capacity)	-101.9	-101.9	-4.393
	(111.0)	(111.0)	(130.4)
Mean Outcome	1225.4	1225.4	1225.4
Year FEs	Yes	Yes	
Plant FEs	Yes	Yes	Yes
State X Year FEs			Yes
N	6263	6263	6263

* $p < 0.1$, ** $p < .05$, *** $p < .01$
Standard errors are clustered at the plant-level.