Title
Three Essays in Applied Microeconomics

Author
Krumholz, Samuel

Publication Date
2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Three Essays in Applied Microeconomics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Samuel Krumholz

Committee in charge:

Professor Josh Graff Zivin, Chair
Professor Prashant Bharadwaj
Professor Gordon Dahl
Professor Zoltan Hajnal
Professor Craig McIntosh
Professor Karthik Muralidharan

2020

The dissertation of Samuel Krumholz is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2020

DEDICATION

I dedicate this dissertation to Akana.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Josh Graff Zivin. He had to put up with listening to more half-baked ideas and reading more unnecessarily long emails than I can count and yet always seemed to know how to distill things to the heart of the matter. This dissertation could not have been written without his wise guidance. I would also like to thank the rest of my committee members, each of whom provided invaluable feedback throughout my Ph.D. career. All of your doors were always open, your comments were always incisive and constructive and you always pushed me to do better. I cannot fully express how much I appreciate the time and thought each of you has taken to help me over the past six years during my time at UCSD.

I next would like to thank my cohort at UCSD and in particular members of the Applied Picnic. From first year in The Commons to the reunion of the Island in Sequoia 225, my classmates made the lowest points of the Ph.D. process bearable and the highest points even better—I cannot imagine undergoing this process without a cohort like the one we were lucky enough to have. While I am very excited to graduate, I'm also sad because I'll be moving away from the lifelong friends made over the past six years. I particularly would like to thank my officemate, friend and co-author Becky Fraenkel. I cannot imagine a better sounding board and look forward to many future collaborations.

I would like to acknowledge and thank my funding sources during the dissertation process. This includes the Sloan Foundation/NBER Fellowship in Energy Economics, which provided two years of funding during my fourth and fifth years (that was particularly fortuitous as it allowed me to live in Boston with my wife during that period), the Charles B. Koch Foundation for Criminal Justice Reform, which provided the funding necessary to obtain the court data used in this dissertation and the UCSD Economics Department, which provided a number of research and summer funding grants that made living on the graduate stipend in San Diego manageable.

I would also like to thank the numerous state and local employees who went above and beyond their jobs to help me with my research. This includes the members of the Washington

Administrative Office of the Courts who answered countless questions about the Washington court data, numerous staff members at state Departments of Education around the country who were willing to dig deep in their archives to provide necessary information about property tax rates and tax bases within their states and hundreds of county clerks who provided me with election results on district attorneys.

I would like to thank my grandparents, parents and siblings who have provided unconditional love and support throughout my life. My parents have always encouraged a love of learning and a deep curiosity about the world and I believe that this, more than anything else, started me on a path that has led me here.

Finally, I would like to thank my wife, Akana Noto, whom I met and married during this journey. She dealt with many lost weekends and nights, was forced to hear more about economics than I'm sure she ever expected to in her life and patiently listened to my practice presentations so many times that I think she could have delivered many of them herself (and likely better than I could). Throughout it all she remained unconditionally supportive, provided essential intellectual feedback, and most importantly, served as a constant reminder that were other, better, things in the world beyond this dissertation.

Chapter 1, in full, is currently being prepared for submission for publication of the material. Fraenkel, Rebecca; Krumholz;Samuel. “Property Taxation as Compensation for Local Exeternalities: Evidence from Large Plants”. The dissertation author was a primary investigator and author of this material.

Chapter 2, in full, is currently under submission for publication of the material. Krumholz;Samuel. “Enforcing Compliance: The Case of Automatic License Suspensions”. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, is currently under submission for publication of the material. Krumholz;Samuel. “The Effectiveness of Litigation as a Policy Instrument: The Case of the New Source Review Litigation”. The dissertation author was the primary investigator and author of this material.

VITA

2009	B. A. in International Relations and Political Science *cum laude*, Tufts University
2014	Masters in Public Policy, University of California, Los Angeles
2020	Ph. D. in Economics, University of California, San Diego

PUBLICATIONS

Adharvyu, Achyuta, Prashant, Bhardwaj, and Samuel Krumholz. "Early Childhood Health and Development in India: A Review of the Evidence and Recommendations for the Future)" India Policy Forum, (2017).

Liu, Gordon G., and Samuel Krumholz. "China's Epidemiological Transition and Health Reform." China's Healthcare System and Reform" in *China's Healthcare System and Reform*, Cambridge University Press. 2017.

Samuel D., David S. Egilman, and Joseph S. Ross."Study of neurontin: titrate to effect, profile of safety (STEPS) trial: a narrative account of a gabapentin seeding trial." Archives of internal medicine 171.12 (2011): 1100-1107.

ABSTRACT OF THE DISSERTATION

Three Essays in Applied Microeconomics

by

Samuel Krumholz

Doctor of Philosophy in Economics

University of California San Diego, 2020

Professor Josh Graff Zivin, Chair

This dissertation is a collection of three essays on applied microeconomics. The first essay examines how local control over property tax revenues generated from large plants affects local jurisdictions' willingness to host such projects. We first demonstrate that property tax payments from plant openings lead to significant increases in local school budgets and that this change is valued by local residents as measured through home prices. We next show that as local jurisdictions become less able to raise and retain property tax revenue from large plants, the number of these plants within the jurisdiction falls significantly relative to nearby jurisdictions that did not experience such a change. These results suggest that increased property tax revenues are an important benefit of large plants and as a result, policies that affect local control over

property taxation can have major unintended consequences.

In the second essay, I examine the effectiveness and equity of automatic driver's license suspensions for nonpayment of criminal fines, a policy that is in place in more than 40 states and that affects millions of drivers annually. Using a unique natural experiment in Washington that first eliminated and then reinstated driver's license suspensions for traffic offense punishment noncompliance. I find that mandating suspensions caused large increases in compliance, fine-repayment, and total punishment with greater effects for lower-income individuals. I further show suggestive evidence that the policy causes declines in traffic accidents among low-income drivers suggesting that such laws are an effective, but highly regressive way to improve traffic safety.

In the third essay, I examine the effects of a major environmental litigation initiative, which led one-third of the US coal-fired power plant fleet to come under a consent decree. I show that legal settlements arising out of this initiative caused large decreases in plant pollution emissions, which further led to meaningful improvements in local air quality and decreases in local cardiovascular and respiratory mortality rates. I conclude by showing suggestive evidence that in regulated electricity markets average electricity retail price and utility revenues increased following a settlement suggesting that a large proportion of the overall costs were borne by ratepayers.

Chapter 1

Property Taxation as Compensation for Local Externalities: Evidence from Large Plants

1.1 Abstract

The external costs and benefits of large capital-intensive projects such as industrial plants, ports and pipelines often occur on dramatically different spatial scales. When local jurisdictions have control over land-use, this spatial mismatch can prevent socially beneficial projects from moving forward or allow socially harmful projects to be built. In this paper, we explore how local control of property taxation, one potentially important localized benefit of these projects, can impact land-use decisions in the context of large plants. We first demonstrate that property tax payments from plant openings are both economically large and valued by local residents as measured through changes in home prices. We next show that limiting local jurisdictions' access to property taxation affects their exposure to large plants by using a series of school finance reforms as plausibly exogenous shocks. Following these reforms, we observe significant declines in large manufacturing establishments and local manufacturing employment per capita both in absolute and relative terms. These results suggest that increased property tax revenues are an important local benefit of large externality-producing projects and that policies which affect local property taxation can have major unintended consequences for non-residential land-use.[1]

1.2 Introduction

When local jurisdictions have control over land-use, proposed projects must create net benefits for the host community in order to be approved. However, many projects create external costs and benefits on vastly different spatial scales. For example, a large manufacturing plant may simultaneously increase local exposure to pollution and contribute to the global risk of climate change, while also boosting regional productivity and employment. This dynamic creates the potential for substantial inefficiencies; communities will refuse to approve projects that decrease

[1]This paper is co-authored with Rebecca Fraenkel. Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at `http://www.zillow.com/ztrax`. The results and opinions are those of the author and do not reflect the position of the Zillow Group.

local welfare even if they benefit society as a whole (e.g. a nuclear power plant that powers a region cleanly, but significantly lowers nearby home values), while approving projects that are locally beneficial, but socially costly (e.g. a plant with large local employment effects that poisons the drinking water of down-river jurisdictions). Because local control of land-use is very common in the United States (Gyourko, Saiz, and Summers, 2008), these types of inefficiencies likely have significant impacts on overall well-being.

In this paper, we study how local government control over the revenues created from property taxation impact these types of land-use decisions. Local property tax payments can act as a transfer from the externality-producing entity to the jurisdictions responsible for land-use. If local governments spend tax payments efficiently, the payments should enhance the value of living within jurisdictions that have these types of projects, increasing their likelihood of approval. Conversely, state and federal policies that constrain the ability of jurisdictions to raise and retain local property tax revenues should mute this effect with potentially large implications for local industrial development and environmental quality.

Our goal in this paper is to evaluate this hypothesis empirically. We begin by testing the extent to which property tax payments from large projects are valued by local homeowners as measured through changes in home prices. This question is important not only as a necessary precursor to the second half of our analysis, but also because these benefits have the potential to change the distribution of groups that gain and lose from the construction of new capital projects; depending on the income and demographic characteristics of individuals inside and outside of the taxing jurisdiction, property tax revenues could either significantly dampen or amplify existing inequities in exposure to the projects' negative local externalities.

Our specific empirical context is the effect of power plant openings on school districts. Power plants exemplify the types of projects that create spatially divergent external costs and benefits,[2] while school districts are the majority recipient of property tax dollars and a major

[2] An additional benefit of power plants from an identification standpoint is that they are relatively free of large positive local externalities such as agglomeration or employment effects.

determinant of home prices and locational choice across the United States. We first look at the effects of these openings on a school district's tax base, property tax rates, revenues and expenditures in order to understand the magnitude of the plant's tax base effect and how this increase in fiscal capacity is used. We then examine the extent to which these changes are capitalized into home values.

To estimate the effects of these openings, we use a border difference-in-differences design, in which we compare outcomes before and after an opening in neighboring districts that did or did not receive a plant. To address concerns about non-random plant siting we also introduce a third element of variation: the expected size of the per-student fiscal impact of the new plant, which we measure by dividing the plant's estimated construction cost by the total number of students in a district in the year of construction.[3] To test the effect of this new tax base on home values, we use Zillow ZTRAX home transaction data and the same border difference-in-differences design, but restrict our sample to only home transactions within a mile of the border between the plant's school district and all neighboring districts to ensure that treated and control homes are similarly exposed to other positive and negative effects of the plant.[4]

We find that property tax payments from these plants are both economically large and highly valued by homeowners. On average, an opening increases a host district's tax base per student by 10%. This tax base increase leads to both a small decrease in property tax rates and a larger increase in educational spending concentrated on capital expenditures. We further find that these changes are valued by local homeowners; home prices increase by 4-5% following a plant opening on the plant's side of the border relative to similar homes directly across the district boundary. These results suggest that property tax payments by large plants act as a substantial local benefit for homeowners within the plant's jurisdiction.

We next test the second half of our hypothesis: restricting jurisdictions' ability to access

[3]We also show robustness to using nearest-neighbor propensity score matching to identify counterfactual districts rather than geographic neighbors, and results are very similar.

[4]Results are robust to a wide range of bandwidths as well as restricting border regions to be greater than 10km from the plant.

property tax revenue should reduce their willingness to be exposed to externality-producing plants. To examine this question empirically, we use a series of state-level school finance reforms over the past half century that dramatically increased the magnitude of state education transfers tied to the level of local property wealth and/or imposed strict property tax limitations, both of which had the effect of reducing the value of tax base increases to local jurisdictions.

To estimate these effects, we employ a geographically-proximate county pair difference-in-differences design in which we compare counties in states that had a reform to nearby counties in states that did not have a reform (or whose reform would occur in the future). Because of data limitations surrounding the timing and location of old power plant openings, we use large manufacturing establishments and manufacturing employment per capita at the county level as our primary outcome of interest.

We show that these reforms led to meaningful (10%) declines in large manufacturing establishments and manufacturing employment per capita in the fifteen years following enactment both in absolute and relative terms. These findings suggests that reducing the tax benefits from large plants has a significant negative impact on local industrial development. Results are robust to a variety of specifications, covariates and weighting schemes and show no evidence of any pre-trends.

This paper makes several contributions to the existing literature. By providing new evidence that property tax payments produced by large externality-producing projects are highly valued by local homeowners, we build on previous work that has estimated other costs and benefits of large plants including decreased health and human capital among individuals affected by pollution (Luechinger, 2014; Barrows, Garg, and Jha, 2018; Persico and Venator, 2018), lower home values near the plant (Davis, 2011; Currie et al., 2015; Gibbons, 2015) and agglomeration and employment benefits (Greenstone, Hornbeck, and Moretti, 2010). Because these benefits accrue to many of the same individuals affected by the plant's negative externalities, they have the potential to act as partial compensation for these costs, which has significant implications for our understanding of income and racial disparities in who is helped and who is harmed by large, essential infrastructure projects (Boer et al., 1997; Banzhaf, Ma, and Timmins, 2019).[5] Finally, by showing that a shock to inputs (non-residential tax base) of public goods provision leads to home price increases, we contribute to a broad public finance literature literature focused on the capitalization of local public goods (Oates, 1969; Black, 1999; Anderson, 2006; Bayer, Ferreira, and McMillan, 2007; Nguyen-Hoang and Yinger, 2011) as well as providing new evidence that local politicians use these tax base shocks for the benefit of local homeowners rather than engaging in capture (Martinez, 2016).

By demonstrating that shifts in local governments' ability to retain property tax revenue significantly affect non-residential land-use decisions and industrial development, we build on a literature examining the development incentives embedded in fiscal decentralization. Previous theoretical work established the importance of local governmental incentives in encouraging economic growth (Weingast, 2009), while empirical work focused largely outside the United States has found support for the idea that local government's share of local (non-property) taxation influences local public good provision and economic development (Han and Kung, 2015; Careaga and Weingast, 2003; Burnes, Neumark, and White, 2011; Zhuravskaya, 2000), but since reforms are often nationwide and come with large income and political consequences, well-identified studies of these effects are scarce (Gadenne and Singhal, 2014). Our results build on this work by presenting novel evidence from a large developed economy that fiscal centralization can have large impacts on local development. This finding is particularly important because in a federal system many higher-level policies aimed at other economic and social goals affect local control over property taxation and our results imply that these policies may have significant unintended consequences.

Finally, these results contribute to a growing literature on the effects of centralizing and

[5] Indeed, in this way, the local fiscal benefits they provide are very similar to those created by natural resource windfalls as shown in Marchand, Weber, et al. (2015), Martinez (2016), Sances and You (2017), and Bartik et al. (2018).

equalizing school finance reforms in the United States. These reforms have been shown to greatly increase low-income students' long-run educational and earnings prospects (Biasi, 2019; Miller, 2018; Lafortune, Rothstein, and Schanzenbach, 2018; Jackson, Johnson, and Persico, 2015; Card and Payne, 2002), while also affecting local housing values by diluting the value of local tax dollars (Hoxby, 2001; Hoxby and Kuziemko, 2004) and changing local property tax burdens (Lutz, 2015; Ross, 2013). Our paper is the first to show an additional major unintended consequence of these reforms—by divorcing the size of the local tax base from available revenue for schools, these reforms affected local non-residential land-use decisions and, in particular, the development of local industry.

The remainder of this paper proceeds as follows. Section 2 provides background on our institutional setting. Sections 3 and 4 describe our empirical strategies and main results. Section 5 concludes.

1.3 Background

In this paper, we first investigate the extent to which local property taxation from large capital projects are valued by local homeowners and then test how limiting this benefit stream affects jurisdictions' willingness to be exposed to these projects. To answer these questions, we undertake a number of separate analyses that rely upon institutional details in plant siting, local public finance, and state school finance systems. In this section, we provide some necessary background information in each of these areas to allow the reader to better understand the validity of the assumptions behind our identification strategies and the plausibility of our observed effects.

1.3.1 Plant Siting

Power plant siting is a complex process governed by a large web of state and local regulations.[6] Utilities take into account a number of factors when siting including access to transportation and energy infrastructure, construction costs and environmental concerns (Cirillo et al., 1977). There is typically a significant trade-off between the low-cost and low-environmental impact of locating in rural areas and increased electricity transmission costs (Davis, 2011).

Utilities also face significant constraints imposed by local, state and federal governments. In general, new plants must be permitted by state and local governments. In 22 states, approval of a specific site does not require approval from the state (although general permits for plant construction are still necessary). In these states, local bodies (typically municipal and/or county governments) have the final say over whether or not a plant can locate in their jurisdiction. Conversely, twenty-eight states have power plant siting boards whose approval is necessary for a plant to locate at a specific site. These regulations appear to have changed little since the 1970s (Cirillo et al., 1977; Ferrey, 2016). In sixteen of these twenty-eight states, the siting board is able to preempt local land-use rules and grant approval to a site over local opposition. In the remaining twelve states, local land-use approval is a prerequisite for siting board approval (although there are some avenues for exceptions). However, even in the sixteen states with preemption powers, local governments are active participants in the permitting decision, and it is unclear in practice how often the wishes of these local governments are overruled. For non-power plants, there are no state siting boards and so local bodies have an even larger say in siting decisions.

Local land-use decisions are typically governed by the local city council (in incorporated areas) or county commissioners (in unincorporated areas). In most states, school districts, the focus of our empirical study, have no control over local land-use.[7] However, in many localities school districts are nearly coterminous with municipalities. For instance, Fischel (2010) finds that

[6] The discussion in this section owes a large debt to Ferrey (2016).

[7] The exception is in New England and in some states in the Mid-Atlantic where schools are run directly by municipalities/the county.

two-thirds of medium-to-large cities in the United States have substantial overlap with a single school district suggesting that municipal or county leaders will internalize any fiscal benefits to the school district. This overlap is likely even larger in rural areas. Further, even if a district is not coterminous with a local zoning jurisdiction, if the harms of a prospective plant within the home municipality/county are concentrated among individuals within the same school district, we would again expect the relevant municipal leaders to internalize their preferences.

1.3.2 Local Taxation of Plants

In almost all states, power plants are required to pay local property taxes. In the majority of states, power plants are assessed by a state body tasked with valuing public utility property, but pay property taxes locally.[8] In a smaller number of states, utility property is both assessed and taxed locally. With few exceptions (i.e. wind power in Kansas), all privately-owned utilities pay local property tax. Taxation of publicly-owned utilities is more complex. Most major publicly-owned utilities including the Tennessee Valley Authority and plants owned by Nebraska's public power districts make payments in-leiu of taxes (PILOT) to local areas. The amount of these PILOTs are typically set by statute and apportioned based on the fraction of a utility's property in a given jurisdiction. Non-power plants are almost always assessed and taxed locally.

Anecdotally, large industrial plants and other projects are recognized to be major contributors to local budgets. In communities nearby plants, local newspapers frequently remark on the magnitude of local power plant tax payments and discuss possible downward reassessments as being disastrous for local communities (Samilton, 2018; Williams, 2018).[9] Public schools receive the majority of property tax revenue and about 40% of state and local education funding on average comes from property taxes. Additionally, local property taxes are often the only source

[8]This can happen either directly or indirectly with the state paying each jurisdiction its share of the total payment based on the proportion of utility property located in its jurisdiction.

[9]Similarly, a large threatened downward reassessment of pipelines in Northern Minnesota was reported as being potentially disastrous for local municipalities and schools.

of funding over which school districts have direct control (Oates and Fischel, 2016).

In Figure A.1 we show the importance of utility and industrial property to the tax bases of districts with plants. The top panel shows the share of total valuation made up by utility property by district generation capacity (100 MW bins) in eight states with local utility valuation data.[10] Among districts with no generating capacity, utility property typically makes up 5% of the total tax base (from infrastructure such as transmission lines and pipelines). However, this proportion rises quickly as generation capacity increases; in districts with 1,000 MW of generating capacity, utilities make up 15% of the local tax base and in districts with 2,500 MW capacity they make up over 30% of the tax base.

In the bottom panel of Figure A.1, we perform similar analysis for industrial plants. The more large polluting plants in a school district (a proxy for exposure to industry), the larger the share of industry as a proportion of total taxable value (bottom panel of Figure A.1).[11] In both of these cases because increases are driven by only a small number of plants, it suggests that these facilities are major contributors to the local tax base.

Property taxes are typically charged as a proportion of the assessed value of local properties. The value of a property upon which taxes may be levied is commonly known as the taxable value and is often some state-set proportion of market value ("assessment-ratio").[12] In some states, utility and industrial property have a different assessment ratio than other types of property leading to a higher or lower effective tax rate. In most states it would be difficult to increase rates on these types of property without equivalently increasing rates on local homeowners.

The process of setting local school property tax rates also differs significantly by state. In some states, tax rates are set annually by local elected officials, while in other states, rates are set by local referendum. Additionally, because many states have created strict limit on tax and revenue growth, meaningful increases in tax rates often must be approved directly by voters

[10]Connecticut, Georgia, Iowa, Minnesota, Ohio, Oklahoma, Oregon and Washington.

[11]We use plants that report to the EPA's Toxic Release Inventory (TRI) as a proxy for large polluting plants

[12]In most states, assessment ratios are created at the state level. In a small minority of states, local control is possible. A notable exception is Pennsylvania, where assessment ratios are set by the county

even if small changes need not be. This is also true for school bonds, which are repaid through increases in the local property tax rates. We discuss this process more in the next subsection.

1.3.3 School Finance Equalization

In response to both court orders and the threat of litigation, many states have undertaken dramatic reforms to their school finance systems over the past fifty years (Jackson, Johnson, and Persico, 2015), moving from primarily locally-financed systems to systems with greater levels of state support. These reforms have typically centered around ensuring some combination of adequacy or equity. Adequacy-based reforms work to ensure that all districts have sufficient funding to provide an "adequate" education to their students. Equity-based reforms work to ensure that large disparities in spending across districts within the state do not exist. In practice, most reforms have some effect on both adequacy and equity. Hoxby (2001) and Jackson, Johnson, and Persico (2014) provide a more extensive overview of the history of school finance reforms in the United States.

Today, most states have a system that at least partially equalizes spending across districts. Although specifics vary from state to state, the vast majority of states have a foundation formula which provides a guaranteed amount of funding for a district based on the number of enrolled students, sometimes weighted by their expected expense of education (i.e. English as a second language students may be worth more than native speakers in the formula). Local districts are then assigned a portion of this formula for which they are responsible ("local share") based off of their local property wealth (or less-commonly a formula including property wealth, income and other determinants of local fiscal capacity). In order to maintain equity, districts in many states also limit the tax rate that districts can charge above the amount that will provide their expected local share (and in some cases, the state can recapture any revenue above a certain threshold). The strictness of these limits varies dramatically across states.

In this paper, we are interested in understanding how changes in the marginal value of a locality's tax base affects industrial development. We study school finance reforms because in many cases they acted as a significant shock to this value both by changing the upper and lower limits on taxes that can be charged and shifting the degree of crowd-out of state-revenues based on property wealth. To see this, note that a simplified funding formula common to many states is:

$$Rev_d = F_d + S_d - \tau^* V_d + \tau_d V_d \tag{1.1}$$

where F_d are federal transfers to the district (which are independent of the local tax base), S_d is the state guaranteed funding to the district, which typically depends on local student characteristics but is independent of local tax base, τ^* is the state-assigned tax rate used to determine a district's local share (typically this is uniform across most districts within a state), V_d is the total assessed value of a district's property and τ_d is the district's chosen tax rate where $\tau_d \in (\underline{\tau}, \bar{\tau})$, state-set limits on the taxes that can be charged. In reality, these formulas are much more complicated, but because we are only interested in the effects of changing tax-base on revenue, this simple illustration acts as a good representation.

All else equal, increasing τ^* implies that growing a district's tax base will lead to a larger reduction in state funding, making any tax base increase less valuable. Similarly, creating a more stringent limitation on the taxes a district can charge also makes new tax base less valuable because districts are unable to fully access the tax base's potential revenue and put it towards their preferred use (assuming that the tax rate constraint binds).

Thus, to understand how a reform affects the marginal value of new tax base, we need to know both how a reform affects $\bar{\tau}$ (through new state limitations) and τ^*. In general, there is no simple summary statistic for either of these terms as funding formulas are written in such a way that the exact level of crowd-out and tax limitations will vary by district. For our analyses, we qualitatively describe these quantities for all states over the past half century using *Public Finance in Public Schools in the United States*, a report issued roughly every five years from 1952 to 2018 that describes the school finance system in use by each state as well as relevant taxation and spending limitations. Using these formula in combination with the narrative descriptions in the report, we then identify major school finance reforms, in which large changes to crowd-out or tax limitations occurred for most districts within a state. Our process for identifying these reforms is discussed in much more detail in Section 4 and Appendix A.4.

1.4 Are Property Tax Payments from Large Capital Projects Valued by Local Homeowners?

We begin by investigating the extent to which property tax payments from large externality-producing projects represent a benefit stream to their local jurisdictions. This question is important for two reasons. One, the value of property tax payments has important implications who bears the cost and benefits of these types of plants—an issue of large policy interest and an area of active debate in the environmental justice literature (Banzhaf, Ma, and Timmins, 2019). Two, in order to answer our second question—how changing jurisdictions' ability to raise and retain property tax revenue affects local land-use decisions—it is first necessary to establish that such revenue is indeed valued by local homeowners. We break our analysis into two parts. We first estimate the magnitude of these payments and show how they are used by local communities. We then estimate how homeowners value these payments as measured through changes in local home values.

1.4.1 Data and Sample Selection

The data to perform the analyses in this section come from four primary sources. First, we obtain power plant location, opening dates, energy source and nameplate capacity from Form EIA-860, published annually by the Energy Information Administration (EIA). We assign each power plant to its 2000 elementary or unified school district using coordinates provided by the EIA and shapefiles produced by the 2000 Census.

Second, we obtain data on taxable value and district property tax rates by collecting information from state Department of Education and Department of Revenue annual reports. This is to our knowledge a novel dataset of longitudinal district tax rates and assessed values in both its geographic and temporal scope.[13] The left panel of Figure A.2 shows our data coverage by state. We have data for forty states and the vast majority of these states have data on both property tax rates and taxable value. The right panel shows coverage over time. By 1999, we have data for over 50% of the districts in our sample and this number increases to over 80% by 2015.[14] It is important to note that states use different assessment ratios (the proportion of the true market value of a property that is taxable) and therefore although rates and assessed values are generally comparable within states over time during our sample period, cross-sectional interstate comparisons of these variables are generally not informative. In our primary analysis, we inflate taxable values (deflate property tax rates) according to reported assessment ratios, however such a conversion is imperfect as states sometimes report a summary taxable value, but have different assessment ratios for different classes of property. As a result, in all regressions we include state by year (or more restrictive) fixed-effects so all comparisons are only made within state-years.[15]

Third, we use district finance, staffing and demographic data from the National Center for Education Statistics (NCES) created by the Rutgers Graduate School of Education Education Law Center (Weber, Srikanth, and Baker, 2016). For all fiscal years between 1995 and 2016, we have detailed data on revenue sources, expenditures by type, district staffing by occupation,

[13]Biasi (2019) and Miller (2018) both collect similar data, but their collection includes fewer states and is over a more limited time period.

[14]Several states fund schools through county or municipal budgets. These state include Massachusetts, Connecticut, New Hampshire, Maryland, Virginia and North Carolina. For these states, we include municipal rather than school district tax rates.

[15]In all analyses involving taxable value or tax rates, we drop New York and Pennsylvania because for most years in our sample, assessment ratios were set by counties and so reinflation is not possible. We also drop Kentucky because reported tax rates were an order of magnitude higher than other states despite the state officially assessing properties at fair market value. This is consistent with anecdotal evidence suggesting county assessment offices in Kentucky systematically undervalue local properties.

student race/ethnicity and free-lunch eligibility. All financial data are inflation-adjusted and presented in 2014 dollars. Our sample consists of elementary and unified school districts that existed in all years between 1995 and 2015, had greater than 200 students and fewer than 50,000 students in 1995[16] and never underwent a boundary change or a district type (i.e. elementary, unified, secondary) change over our 22 year time-period.[17] This leaves us with 11,824 total school districts.

Fourth, we use home transaction data from the Zillow ZTRAX database. This nearly-nationwide database contains almost all home transactions between 2005 and 2017 with much longer temporal coverage for some counties. There are 12 states for which home transaction data are not publicly available. Data include information on sales price, home attributes, home location and owner characteristics. Appendix A.2 provides greater detail on how the ZTRAX data were processed for this project. We currently use home price data from 13 states that have both comprehensive transaction coverage and a large number of plant openings. In the future, we plan to expand our analysis to all states with available data.

1.4.2 Effects on School District Budgets and Property Taxes

Empirical Strategy

We begin by estimating the effect of power plant openings on school district fiscal outcomes. We focus on power plants because although they create significant negative local externalities, they do not cause many of the positive local externalities such as large employment or agglomeration effects created by other types of capitally-intensive projects. This simplifies our empirical efforts to estimate the property taxation benefits provided by these plants because it

[16]Because our primary outcome variables are per student, we want to exclude very small districts where small changes in the student population could lead to large changes in the outcome variable. Because most metro areas typically have a single large center-city school district, there is a concern that large districts are unlikely to be a good counterfactual for neighboring districts.

[17]We exclude districts that have undergone boundary changes to ensure that any observed changes are not simply arising from changes in composition within the district.

is less plausible that any observed effects are being driven simply by greater economic activity within the district.

In an ideal world, we would randomly assign plants to some districts, but not others and examine the changes in fiscal outcomes that ensued. Instead, we employ two separate but related strategies to approximate such a randomization. We first use a border difference-in-differences design in which we compare tax base and school finance outcomes in districts experiencing an opening between 1995 and 2015 relative to neighboring districts that never experience an opening during this time period. We use neighboring districts as controls because they are exposed to many of the same local economic and school funding shocks as treated districts and so plausibly would be expected to have similar outcomes to the treated districts were they not to have received a plant—in subsequent sections, we provide extensive evidence supporting this assumption.

We define a treated district as a district receiving a new utility-owned natural gas or wind turbine in any year between 1995 and 2015—our sample is made up of 55% natural gas plants and 45% wind plants and 10% solar plants.[18] In this paper, we are interested in the property tax effects of large projects—accordingly, we only include plants that are above 25MW in size, a common cut-off used by the EPA when determining eligibility for pollution control regulations.[19][20] There are 1,297 such plant openings in our data.[21]

We define the treatment year as the year in which the plant first obtained regulatory approval or began construction because knowledge that a jurisdiction will receive a plant in the future may affect current taxing or borrowing behavior. For the 25% of plants for which this

[18]These are the plant types for which we can reliably estimate construction cost, a necessary component of one aspect of our identification strategy. They also make up the vast majority of new openings during this time period.

[19]i.e. only coal units greater than 25MW are required to participate in the Continuous Emissions Monitoring System (CEMS))

[20]For reference, if running 100% of the time a 25MW plant could provide power for roughly 25,000 homes. Since most plant's capacity factor is far below 100%, these plants provide power for closer to 12,000-24,000 homes depending on the plant type.The vast majority of plant openings smaller than this level are very small (<10 MW) solar or landfill gas installations.

[21]We additionally drop 50 plants whose first year reporting to the EIA is more than 2 years after their stated operation date, 3 plants whose construction date is after their stated operation date and 29 plants whose operation date was more than 5 years after construction approval.

information is not available, we instead use the year of operation—note that if districts begin borrowing or raising revenue after approval, using year of approval would bias our results towards zero. Because our outcomes are at the district level, we restrict our analysis to only the first plant opening in a district over our timeframe. In total 852 districts experienced an opening during this time period.[22] Of these districts, 675 fit our district sample criteria[23]

In the primary analysis, we restrict our sample to only those plants whose approval year was between 2001 and 2007 in order to create a fully-balanced panel. This addresses concerns described by Goodman-Bacon (2018) that a difference-in-differences model with time-varying treatment and an unbalanced panel can place a negative weight on some years, producing an estimand of different magnitude (and potentially sign) from the average treatment effect. However, we also show in robustness checks that using the full unbalanced panel produces very similar results.

For each of these districts we identified all neighboring districts within the same state that did not also experience an opening and fit our sample inclusion criteria (i.e. in existence over the full period with greater than 200 students or fewer than 50,000 students at baseline). If control districts border multiple treatment districts then they will enter the sample multiple times. To adjust for this, we use two-way clustering at the plant district (all bordering control districts attached to a treated district) and district level. Different districts have different numbers of border pairs and so an unweighted regression would overweight districts with high numbers of pairs. To adjust for this, we weight all districts attached to a given opening by the inverse of the total number of border pairs the opening has in a given sample year. Using this sample, we then implement the following difference-in-differences specification:

[22] 75% of these districts only experienced a single opening. Results are robust to restricting our sample to only these districts.

[23] 59 districts had fewer than 200 students in 1995, 16 districts had greater than 50,000 students in 1995, 81 districts were not in operation for all 22 years or experienced a boundary change, 3 districts had fewer than 3 grades, 2 districts had greater than 1 log point annual change in students.

$$Y_{dpt} = \alpha_d + \tau_{pt} + \beta Post_{dpt} * Treat_d + \varepsilon_{dpt} \tag{1.2}$$

where Y_{dpt} is the outcome variable in district d, in border pair p and time t, α_d is a district fixed-effect, τ_{pt} is a pair by year fixed effect, $Post_{dpt} * Treat_d$ is our variable of interest, the interaction between the period following a pair's plant beginning construction and whether or not a district is the treated member of the pair, and ε_{dpt} is a mean-zero error term. Because our treatment year is defined as the year of approval, we separately examine effects for two post periods (years 0-2 which are the years over which most plants are built and years 3-8 when most plants are already in operation). In our primary analysis, we show results for both all openings and "non-small" openings (defined as having fewer than \$10,000 per student in expected tax base increase based on plant and district size—$\approx$ 10% of sample), as these small openings are unlikely to produce sufficient tax revenue per student to make a meaningful impact on district finances.

Table A.1 shows differences in baseline covariates between treated and control districts. Unsurprisingly, given previous work in the environmental justice literature, there is a significant 2.9 pp difference in the proportion of underrepresented minorities (Black and Hispanic students) in the treated district relative to the bordering controls. However, reassuringly, there are no economically or statistically significant differences between treatment and control districts in number of students in a district, student free and reduced lunch status, school revenues or local home values. Of course baseline differences do not invalidate our design—what matters for identification is parallel trends—but nonetheless the fact that differences across most key covariates are minimal increases confidence in the validity of this strategy.

There are two major related challenges to identification in this setting. First, it is possible that treatment districts are fundamentally different from control districts and so even in the absence of treatment they would be expected to have different trajectories in the outcome variables. We address this possibility in several ways. First, we examine whether pre-trends exist in major demographic and economic variables that we might expect to be correlated with both plant openings and changes in tax base—there is no evidence for any such trends. Second, we estimate dynamic versions of all models and examine trends prior to plants receiving regulatory approval for construction—we again find no evidence for any violations. Third, we show that results are robust to a number of different weightings, sample selection criteria and covariates,

Finally, we estimate an alternate identification strategies as a robustness check. First, we leverage the fact that the expected revenue per student effect of a plant opening in a district is a function of both the size of the plant's estimated effect on the tax base and the number of students that this new revenue will be split across. Accordingly, we estimate the expected tax base impact of an opening by dividing a plant's estimated construction costs (a proxy for the plant's value)[24] by the total number of students enrolled in the district in the year the plant received regulatory approval and include this variable in a triple difference framework. The underlying assumption of this analysis is that plants that have a higher value per student should create correspondingly larger tax base effects in their host district. For omitted variable bias to exist in this approach it cannot only be the case that receiving districts are systematically different from their control districts in ways that are correlated with time-trends in the outcome variables, rather it must instead be the case that differences between treatment districts receiving openings with larger expected fiscal impacts and their assigned control district are systematically different than the differences between treatment districts receiving openings with smaller expected fiscal impacts and their assigned control districts.[25]

A second and more challenging barrier to identification is the possibility that a jurisdiction's decision to open a plant is correlated with other factors that may be associated with our

[24]The vast majority of states use original construction cost as the only or primary method of assessment. States that rely on other methods typically use either fair market value (which will be correlated with construction costs) or total production or income, both of which should be correlated with construction costs. Estimated construction costs are based off of fuel and prime-mover specific estimates of overnight construction costs per megawatt-hour in the EIA's Annual Energy Outlook from 1997-2018. More details on these calculations are provided in Appendix B.

[25]Note that this is not simply the difference between districts receiving large and small plants, but is instead the interaction between plant value with the size of the receiving district. A small plant in a small district may have a similar expected fiscal impact as a large plant in a large district because the increased taxable value is split across fewer students.

outcome variables. Under this scenario, it could be true that control and treatment districts are ex-ante similar, but some event (i.e. the election of a developmentally-minded mayor) leads to both the construction of a plant and other changes correlated with an increased tax base or education spending, which would lead us to estimate the plant's effects with bias. Incorporating treatment intensity into our model helps address this concern, but does not fully solve it—it could be the case for instance, that the districts that receive the largest fiscal impacts have the most developmentally-minded leaders. Thus as a second test, we examine if the plant opening is correlated with openings of other types of (non-utility) environmentally harmful plants using data from the EPA's Toxic Release Inventory.[26] If local governments are attempting to attract new facilities, we would expect to see such an increase. Figure A.3 show the main results. There are no significant spikes in openings of non-utility toxic facilities following (or prior to) the beginning of the start of plant construction. This provides at least suggestive evidence that the construction of a plant is not a proxy for a larger development boom.

Results

Table 1.1 show the main effects of a plant opening on major tax base and school finance outcomes using our primary specification. A new plant increases the local tax base per student by 11% on average suggesting that plant openings can have large effects on the fiscal capabilities of local districts. This increase is used primarily to increase local school revenues; there is no economically or statistically significant effects on the local property tax rate. Conversely, a plant opening increases locally-raised district revenues by $500/student (10%) and total revenues per student by $380. This gap is explained by a reduction in state funding. Many state school funding formulas tie the level of state transfers to a district's property wealth—increases in the property wealth should leads mechanically to lower levels of state transfers for education and that is precisely what we see here. Finally, Column 5 shows that expenditures increase by $770/student

[26] Although imperfect, the TRI provides the best publicly-available record of new plant openings. We say a plant has "opened" if it is the first year in which it appeared in the TRI.

or more than 5% following an opening. These results demonstrate that having a plant enter a school district has important effects on local educational spending.

Figures 1.1 and 1.2 show these results in event study form. Reassuringly, across all outcomes there are no trends in the six years prior to plant approval. We then observe a rapid increase in tax base per student, local revenues per student, total revenues per student and total expenditures per student beginning two to three years following plant approval, after which these outcomes plateau at a significantly higher level. The opposite pattern is true for property tax rates, although the decrease is not statistically significant. This timeline makes sense intuitively; construction was typically completed two to three years after final approval and so this is precisely when we would expect to see the change in tax base (and by extension district revenues) appear in the data.

Table 1.3 shows where the additional revenues created by the plant are spent. We first see that openings lead districts to take a large amount of additional debt ($750-$1,100 per student or 15%-25% increase). This increase in debt explains why plant openings appear to have a larger effect on expenditures than on revenues. This debt increase may occur for two reasons. One, because the plant opening increased the size of the tax base the price in additional tax rate increments for any given sum has now fallen for a given household. Two, in many school funding formulas, debt allows local governments to use their additional tax base to increase school budgets when any other increase in local revenue would simply crowd out of state transfers.

In general, most school district debt is used to fund capital expenditures. We observe a similar phenomenon here. Despite making up only 10% of total school spending, the majority of expenditure increases caused by the plant occur on capital projects. Specifically, by 4-10 years after approval, spending on capital projects increases by $400/student (33% increase). There is also evidence for smaller increases in non-instructional spending and instructional salaries. The disproportionate use of new revenue to fund capital expenditures is consistent with previous work examining school district responses to other forms of revenue shocks (i.e. (Davis and Ferreira, 2017)).

We finally examine changes in school district revenues by source. If these changes are driven by plant openings, we would expect several dynamics to hold. First, changes in local revenue should be driven through property tax revenue, increased parent government contributions for districts that are not financially independent[27], or payment-in-leiu of taxation (PILOT)[28] We should expect no change (or a compensatory decrease) in other sources of local funding such as sales and income taxes. Second, we should expect a decrease in state aid from the state's school finance funding formula as almost all states now tie formula aid to a district's level of property wealth. Finally, we should no expect no changes in other state funding or federal funding.

Table 1.2 shows the main results of this analysis and they accord exactly with the predictions above. The bulk of the increase in local revenues comes from increased property tax revenues. However, we also see significant increases in parent government contributions and income from unspecified sources, which likely comes from PILOT payments. There is no chance in local revenue from other sources, which largely consists of sales and income tax revenue. We also observe a significant decrease in state transfers through formula aid; this is precisely what we would expect given that formula aid in most states is inversely proportional to a district's level of property wealth. Finally, just as we would expect there is no change in other sources of state funding or federal revenue. These results provide additional suggestive evidence that the observed changes are indeed driven by the plant opening and not other correlated trends in district school finance.

We next turn to examining the robustness of our results to different specification and sample choices. Table A.2 shows results after controlling for baseline covariate by state by year

[27]In some states (primarily states in the Northeast and Mid Atlantic) as well as some large cities), school districts are funded through municipal budgets rather than having independent budgetary and taxing authority. In these states, increased local property tax revenue would be classified as parental government contributions rather than property tax revenue since the district's revenue is technically coming from the parent government.

[28]Some plants may negotiate tax abatements with local school districts that lead them to make transfers to local district outside of the tax system. Further, municipal, state and federal utilities are often mandated by law to pay PILOT because they are exempted from local property taxation.

fixed-effects. If baseline differences between treated and control districts were driving results, we might expect to see effects diminish once these trends are controlled for. However, results remain nearly identical to the full specification. Table A.4 shows effects when including all districts instead of excluding those with a very small expected tax base per student impact. Unsurprisingly, effect sizes fall slightly but remain highly economically and statistically significant. Table A.3 shows results with an unbalanced panel including ten years of data before and after an opening. Results again remain largely unchanged. Lastly, Table A.5 shows effects separately for natural gas openings and wind turbine openings. In both cases, results are highly economically and statistically significant suggesting that effects are not driven by one type of opening. Effects are about twice as large for wind turbine openings than natural gas openings—this is likely because wind turbines open in largely rural areas with little property wealth per student where the expected tax base effect per student is much larger.

Table A.6 shows the results of our triple-difference design. We use the same sample as our primary analysis, but interact the post by treat term with logged expected tax base per student. All outcome variables are also logged to create an elasticity[29]. Results are highly consistent with the primary analysis despite being identified off of an entirely different source of variation (opening vs no opening relative to the slope based on size of opening). A 10% increase in expected tax base per student leads to a .4% increase in tax base per student, a .6% increase in local revenue per student, a .13% increase in total revenue per student and a .2% increase in expenditures for student. For context, moving from an opening at the 25th percentile of expected tax base per student to an opening at the 75th percentile results in an increase of 2 log points, or a 8% increase in tax base per student, a 12% increase in local revenue, a 2.5% increase in local revenue and a 4% increase in total expenditures.

Interestingly, one exception appears to be the effect on property tax rates. These rates appear to increase with expected tax base size—a 10% increase in tax base leads to a 2.6 mill

[29]We log expected tax base per student because this variable is heavily right skewed and results are otherwise sensitive to how we treat and define outliers.

increase in tax rate. Likely, this is because increased tax base has two competing effects on the tax rate. An increased tax base makes district's richer, which through an income effect should decrease rates. However, at the same time an increased tax rate lowers the price of debt (which is financed through tax rate increases to repay bonds), which can lead to tax rate increases. In this case, it appears that larger tax base increases are more likely to induce districts to undertake large capital projects and therefore higher rates.

As a final robustness check, Figure 1.3 examines whether these changes lead to large demographic shifts in the composition of students as well as whether there were any pre-trends in these variables prior to the start of construction. We focus on black and Hispanic share of enrollment, the proportion of students eligible for free lunch (a proxy for poverty) and log total enrollment. There are no significant trends in any outcome prior to plant approval increasing confidence that any changes in school finance outcomes were not driven by underlying trends in district demographics in treatment relative to control districts. We also do not observe any large changes after approvals—there is some evidence for a slight decline in free lunch students seven to eight years after approval, which suggests some sorting may be beginning to occur in the long-run, but this effect is not statistically significant. Together these results suggest that our observed effects are unlikely to be driven by differential trends across treatment and control districts nor are driven by sorting occurring after a plant opens

In sum, these results suggest that plant openings can have large impacts on the finances of host school districts. Through their direct tax base effect, these plants lead to large increases in local revenue per student. Despite some crowd-out of state-transfers, openings The increased tax base caused by plant openings combined with the structure of state school finance formulas also induce local districts to take out more long-term debt and finance capital projects, leading to an even larger increase in expenditures per student with no changes in local property tax rates.

1.4.3 Effect of Plant Openings on Home Prices: Empirical Strategy

The previous section showed that property tax payments arising from plant openings lead to appreciable increases in educational spending centered on capital expenditures. However, it is unclear if such spending actually increases the well-being of local homeowners. Instead, such payments could be captured by local bureaucrats or simply spent in a way that was not valued by homeowners. Accordingly, we now turn to the second part of our question: estimating if the fiscal benefits created by plant openings are actually valued by local residents as measured through changes in home values.

Empirical Strategy

In this section, we want to estimate the hedonic value of the increase in tax base caused by a plant opening. To obtain a valid estimate of the effect of tax base increases on home values, we need to compare homes whose values are both expected to evolve similarly in the absence of a plant opening *and* are similarly affected by the non-fiscal positive and negative effects of the opening. In other words, we want to hold exposure to all other positive and negative plant effects constant and just estimate the home price effect of the tax base increase. Thus, our results from this section should not be interpreted as the net home price effect of a plant opening, but instead as the component of the net effect caused by its effect on local school district fiscal capacity.

As described above, one advantage of using power plant openings as our setting is that these plants have relatively small agglomeration and employment effects and so we are mostly concerned with differential exposure to these plant's negative externalities. Accordingly, to estimate the home price effects of the property tax shock alone, we use the same border difference-in-differences approach as above, but instead of comparing whole neighboring districts, we restrict our sample to only a narrow bandwidth around the border. Specifically, we create border pairs between all homes that are in bordering districts with a plant opening and neighbors with no openings and are within two kilometers of the border. We then compare the relative change in housing values on either side of the opening after plant approval.[30] Both sets of homes should be exposed to similar economic and pollution shocks from the plant, but only homes on the plant's district's side of the border will get the benefit of the expanded tax base—if the parallel trends assumption holds, we can then attribute any changes in home prices to the fiscal effect of the plant. We perform this analysis for openings in fourteen states that have comprehensive home price data and the largest number of openings.[31]

One important caveat is that many school district boundaries are shared (or nearby) county and municipal boundaries. This implies that results should not be interpreted as the home price effect of the increased school district tax base alone, but as a weighted estimate of the increased tax base across all local government units that share the same border. We test for robustness by excluding district boundaries that are shared with county boundaries and results are qualitatively similar, but there are many other local government taxing units (i.e. municipalities, irrigation districts, sewage districts, etc) which we lack granular enough geographic data to exclude. Further, even if we could exclude these districts, our remaining sample of school district boundaries would likely be too small to obtain valid statistical estimates. Because of these shared boundaries, we do not use the triple-differences approach above as our primary specification. Home values will respond to the expected fiscal impact on all relevant governmental units and while we expect homes inside districts with large expected valuations per student to experience larger home price increases, there is no reason to believe there is necessarily a monotonic relationship between the expected impact on a district and the expected impact on the district's county or other governmental units (i.e. a plant opening in a small district could be in a very large county or city, while a plant opening in a moderately sized district may be in a very small municipality) making estimates hard to interpret. However, for completeness, we also include the results of this analysis and they are broadly similar.

[30] If a home is near multiple borders we assign it to its nearest border.

[31] Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Our primary specification is as follows:

$$Y_{idpt} = \alpha_d + \tau_{pt} + \beta Post_{dpt} * Treat_d + \varepsilon_{dpt} \tag{1.3}$$

where Y_{idpt} is the logged sales price of home i in district d in border pair p and year t. The vector α_d contains indicators for extremely granular spatial controls. In the primary specification we use border pair x district x .004 degree x .004 degree latitude and longitude cells to ensure that we are comparing homes in very similar neighborhoods, but show robustness to more or less granular controls including a parcel indicator in which our estimand is completely identified off of multiple sales of the same parcel. The vector τ_{pt} contains indicators for border pair by year by month fixed-effects to control for any time-varying characteristics of homes within the border region. The coefficient on $Post_{dpt} * Treat_d$ is our coefficient of interest. Our primary model uses a bandwidth of 2,000 meters, but we show robustness to alternate bandwidths (1,000 meters-3,000 meters).

As with the analysis above, we weight observations such that each plant-opening year counts equally. One challenge with that approach in this setting is that some openings in rural districts have very few (i.e. < 2) transactions per year. Because a single transaction per year provides a very noisy estimate of how local home values are changing, using these openings in our analysis would decrease our power considerably. As such, we restrict our sample to only bordering districts that each have at least 10,000 population, while showing robustness to including all openings. As above, we restrict our primary sample to only openings that have an expected tax base impact larger than $10,000 per student.

An alternate approach is to run the regression at the home transaction level. This implicitly gives more weight to openings whose border regions have a higher number of transactions. The challenge here is that transaction dense areas typically are part of school districts with large numbers of students—as such, these are exactly the openings where we expect the fiscal impact per student (and by extension, the effect on home prices) to be lowest.

The Zillow home transaction data used here does not go back as far as the school finance data for many counties. As a result, we cannot use a fully balanced panel here as we will have too few openings to have sufficient power to detect reasonably-sized effects. Thus, our primary specification uses an unbalanced panel in which all openings have data at least two years prior and two years subsequent to the plant approval year. We also show robustness to using a balanced panel.

The primary threats to identification using this empirical approach are twofold. First, as with the school finance analysis, homes outside the district may not be good counterfactuals for homes inside the district. We test this assumption in several ways. First we employ a dynamic difference-in-difference analysis to test for pre-trends and find no evidence for any violations. Second, we control for a large number of major hedonic characteristics to ensure that we are comparing similar homes in both the treatment and control districts and find results are similar. Third, we show that effects are completely driven by plants with larger expected tax base impacts; there is a minimal and insignificant effect among districts with a small expected tax base impact. This result suggests that there is nothing specific about the types of districts receiving plants that are driving our results.

An additional threat to identification is that the border design does not fully control for other positive or negative effects from the plant. For instance, if the plant increases nearby housing demand, this may increase home values closer to the plants, which will disproportionately be inside the plant district's border region. As power plants do not typically create large amount of jobs or have large agglomeration effects, these types of positive effects are less of a concern. Instead, it is more likely that as the bandwidth of included sales increases, there will be unequal exposure to the plant's pollution, which would likely bias our results downnward. We test for this possibility in two ways. First, we show robustness to a large number of border bandwidths and results are qualitatively similar. Second, we show results when including only border regions that are "far" from the plant (where far is defined as the nearest home to the plant being at least 10km away) as these border regions should be "uncontaminated" by other effects of the plant. Again results do not exhibit large changes.

A final identification check leverages the existence of California in our data. California's school funding formula is very strongly tied to a localities property wealth; outside of taking on debt, it is almost impossible for a school district to increase its revenue flow from an increase in the tax base because any revenue increase will be crowded-out one-for-one by a decrease in state funding.[32] Thus, in addition to showing the pooled results, we show results separately for our California and non-California samples. If the observed home price effects are truly driven by the local fiscal effects caused by a plant opening, we would expect that the results should be larger for the non-California states—indeed that is precisely what we see in the data.

Before moving to the results, it is important to reiterate here that this specification attempts to estimate the effect of the tax base increase on home prices alone, *not* the net effect of a plant opening. Previous research has shown both wind turbines and natural gas plants have negative effects on the values of nearby homes. Because our goal here is to understand the extent to which the tax base increase from these plants is valued by local homeowners, our aim is to hold these negative distance-based effects constant and estimate only the effect of the tax base increase itself. Observing positive effects here does not contradict these earlier results—they are estimating something akin to the slope of home price changes with respect to distance to the plant, while we are estimating how the intercept of being in a given district that receives a plant changes after the plant enters. Both effects are possible and their relative magnitudes will dictate both the average net effect of the plant and who gains and who loses from a plant entering.

[32]Technically, the formula is even more extreme. Most school districts in California do not receive their base local funding from the tax base in their district, but rather as a share of the total county tax base. The exception to this is taxes raised to repay bonds, which are levied only on the tax base within a district's borders.

1.4.4 Effect of Plant Openings on Home Prices

We begin with an estimate of the home price capitalization of the increased tax base created by plant openings. Table 1.4 shows the main effects. Column (1) show the results for our full geographic sample when restricting to districts with greater than 10,000 population. By 3-8 years following approval, home prices increase by 4.8% suggesting that homeowners meaningfully value the increased tax base created by the plant. In Columns (2) and (3) we can see that as expected these results are much larger when excluding California, which has a very strong school finance equalization system, from our sample. Home prices increase by 7% outside of California, but only by 2.8% within California and the effect is not statistically significant. In Columns (4)-(6), we show the same specifications, but including all districts, even those with very low populations. Results remain very similar.

The left panel in Figure 1.4 show these results in event study form for our primary specification. Because we are only identified off of 71 openings, individual year estimates are somewhat imprecise, but nevertheless, there do not appear to be any trends prior to plant approval and then a sustained increase which plateaus three to four years after approval. This provides some evidence that the observed effects are not simply driven by differential trends between the border regions. The right panel in Figure 1.4 shows the home price effects across a number of different distance bandwidths. Coefficients are extremely similar for any bandwidth between 1 kilometer and 3 kilometers suggesting that our results are not driven by bandwidth choice.

Table 1.5 shows results using different expected tax base cut-offs. Columns (1)-(3) show results weighting all plant opening years equally, while Columns (4)-(6) show these same results, but weighting by transaction, which implicitly gives more weight to more populous districts. Several trends are immediately apparent. First, even when including openings with less than $10,000/student in expected tax base, results in the weighted specification remain economically and statistically significant. Second, as we include a higher expected tax base cut-off, effect sizes increase significantly. For districts with an expected increase greater than $75,000/student, home prices increase by 7% and for those with an expected increase greater than $150,000/student, home prices increase by 10.5% in the weighted specification and by 5% and 7% in the unweighted specification where the average expected tax base effect is lower because high student districts receive more weight. This is precisely what we would expect if effects are indeed driven by the tax base effect of a plant and not other, underlying district characteristics. Finally, effects disappear when we include the full sample in the unweighted specification because the average expected tax base per student shrinks significant ($38,000/student compared to $108,000/student in the weighted specification). Again, this is consistent with effects being driven by the size of the fiscal impact.

One potential confounder of our results would be a change in the composition of homes in treated districts after the plant opens. We attempt to control for this in our main specification by controlling for hedonic variables and very granular geographic fixed-effects, but it is of course possible that other, unobserved home characteristics are driving the results. We address this concern in two ways. First, Table A.7 tests whether key hedonic characteristics differentially change in homes sold in treated districts after the plant receives approval. There are no economically or statistically significant differences in lot size, home age, square footage or whether or not a home is single-family home. There is a marginally significant decrease in the number of bedrooms in homes sold, but this effects is very small (x) and would appear to suggest that homes sold after the plant opening are if anything lower value. Second, Table A.8 shows results using our main specification, but including only repeat sales. Standard errors increase as our estimates are only identified off of homes that have multiple sales within our sample period, but results are very similar to our primary analysis suggesting that compositional changes in the underlying types of houses sold are unlikely to be driving our results.

We can also examine if the quantities of houses sold changes after the opening. Because the plant opening causes a large shock to local public good provision, we might expect that households will respond by re-optimizing, increasing home sales, Further, if, as we showed above, the plant causes the tax price of public good provision to fall and there are few supply constraints, we may also expect that districts receiving a plant will see an increase in new construction (Lutz, 2015). Note that as long as the composition of homes is not changing conditional on our covariates (which includes controls for new homes), an increase in quantity will not bias our home price estimates. Table A.9 presents some evidence that the opening does indeed induce re-optimization; home sales in plant's district increase by 22 sales (25%) following an opening, although the effect is only marginally significant. This increase is driven by both sales of old homes and new construction—the probability that a plant district has any new construction increase by 8 percentage points or 20% suggesting that builders are responding to the decreased tax price of public goods created by the plant pening. To ensure that our price results are not driven by new construction, Table A.10 shows our main specification, but excluding newly constructed homes. Results are if anything larger with new homes excluded from the sample.

Table A.11 shows a number of additional robustness tests that aim to rule out alternative explanations for these effects. Column (1) shows results when restricting to border pairs whose closest house to the newly-opened plant is at least 5km away, while Column (2) shows results when dropping boundaries that are also county borders. If anything, home price effects are larger with both restrictions suggesting results are unlikely to be driven by other direct effects from the plant (i.e. land payments or increased housing demand), nor by increases in county fiscal capacity. Column (3) shows results using a fully-balanced panel. Standard errors increase as our sample size shrinks, but effect sizes remain similar suggesting that our unbalanced panel is not driving results. Columns (4)-(7) show results with different levels of time and geographic fixed-effects. Regardless of whether we use district fixed-effects, .004 degree, or .001 degree as our geographic region fixed-effects, results remain extremely similar. The same is true whether we use border by year by month fixed effects as in our main specification, or border by year fixed-effects. Together, these results provide additional reassurance that the observed home price effects are not driven by a single specification choice.

Table A.12 shows results by plant type. As with the school finance results, effects are significant for both natural gas plants and wind farms, but much larger for wind farms. Districts receiving a natural gas opening have to a 3% increase in home prices, while those receiving wind turbines have a 12% increase. As above, this difference between the two is likely because wind farms are disproportionately located in rural areas with few students so they lead to larger increases in tax base per student.

Finally, Table A.13 shows the effects of the opening on school finance variables for the subset of openings studied here using the same border difference-in-differences specification as in the housing price regression. The odd columns show results for the full sample and the even columns show results without California. When including California, results are sightly smaller in magnitude than were observed in Section 3.1 ($\approx$ \$300/student increase in local revenue/student, \$250/student increase in total expenditures per student) and statistically insignificant as standard errors increase relative to 3.1 given our smaller sample size. When excluding California, effect sizes more than double and become statistically significant reflecting California's strong school funding equalization program. The fact that we see correspondingly larger home price effects when excluding California from our analysis adds confidence that we are indeed uncovering the causal effect of the effect sizes are now larger and statistically significant for total revenue.

Together, the results in this section imply that the increased tax base caused by new plants is being used in ways that are valued by local and prospective homeowners. Such a result does not necessarily follow from more local spending; given rational voter inattention it is certainly plausible that local bureaucrats could capture this additional revenue through higher salaries or wasteful spending that bring no benefit to local homeowners, but this does not appear to be occurring in practice. One mechanism through which this capitalization likely occurs is the construction of new schools and other capital improvements. We lack exogenous variation on school construction conditional on receiving a plant, but previous work has suggested that school construction leads to a roughly 6% increase in home prices, even when it is funded using increased taxes on existing properties (Cellini, Ferreira, and Rothstein, 2010; Lafortune and Schonholzer, 2019). This effect size is very similar in magnitude to our observed results.

1.4.5 Valuation of Negative Externalities

The above analyses provide evidence that the tax base increase caused by an entrance of a plant leads to a meaningful increase in school district home values, all else held equal. But of course all else is not held equal—the plant opening also brings with it significant negative externalities. In order to better understand the relevance of the positive effects identified above to homeowners, we attempt to benchmark these effects by estimating he effect of these negative externalities on nearby residents.

We use a spatial difference-in-difference model comparing homes that are nearby the plant relative to those that are further away before and after a plant opens. In particular, our design follows previous work by Davis (2011) and Gibbons (2015) who estimated the negative effect of natural gas plant openings and wind turbines respectively on home prices.The main innovation in our analysis is that we include school district by year fixed-effects in order to control for any positive fiscal effects the plants may have on home values. In this way, we are ensured of estimating the negative effect of the plant's non-fiscal externalities alone and can hold constant the benefits of the district's increased tax base. Specifically, we estimate:

$$Y_{idpt} = \alpha_{dpt} + Dist_{idp} * Post_{dpt} + Z_{idpt} + \varepsilon_{idpt} \tag{1.4}$$

where Y_{idpt} is the log sales price of home i in district d near plant p in month-year t, α_{dpt} is a plant by year by month fixed-effect, $Dist_{idp}$ is a variable capturing distance to the plant (either 5 kilometer bins or log distance), $Post_{dpt}$ is a vector of indicators for various time periods following the plant opening, Z_{idpt} is a .004 latitude x .004 longitude fixed effect, which compares homes within the same .25 mile by .25 mile grid cells and ε_{idpt} is a mean-zero error term. We restrict our sample to only plants that are larger than 100 megawatts in order to ensure that we are considering plants large enough to have a community impact; we also perform a robustness test in which we interact our post by treat indicator with plant size under the assumption that effects should increase with the size of the plant. As before, we weight all plant openings equally and therefore restrict our analysis to only plant openings in districts with greater than 10,000 population. Our primary analysis considers homes within 20 kilometers of the plant, but we show robustness to other cut-offs. We also exclude homes within 500 meters of the plant because there is a concern that these parcels could have purchased in the construction of the plant themselves, which may bias our results.

Table 1.6 shows the main results. Columns (1) and (2) show the results with indicators for home distance to plant for an unweighted and weighted regression respectively. Relative to homes 10km-20km away, home prices fall by 4%-6% within 5km of the plant. Columns (3) and (4) show that this result is not simply a function of the distance bins used; here we interact log distance with plant openings and show that for each log point closer to a plant a house is, its price falls by 3-4% following an opening. Finally, Columns (5)-(8) provide an additional robustness check by interacting our post x distance variable with the size of the plant—if the effect is truly driven by the plant opening, we would expect the negative effect to be larger for bigger plants. Indeed, that is precisely what we see here; relative to homes 10km-20km from the plant, for each 100 MW increase in plant size, the home price effect of the plant decreases by an additional 1.5%. We see a similar result when using log distance instead of distance bins.

Figure 1.5 shows these results in event study form with the log distance specification on the left and the distance bin specification on the right. In both cases there appear to be no trend prior to approval and then a sharp decreases in prices for homes closer to the plant[33] beginning after two to three years approval, which that plateaus at the new lower level. This is precisely the time pattern we would expect if the price change was caused by the plant opening. Table

[33]Note that for log distance this appears as a positive coefficient because it is the effect of being an additional log point away from the plant site.

A.14 tests the robustness of our main specification to different sample restrictions and covariate inclusions. Columns (1)-(2) show the results when restricting to homes within 15km of the plant instead of 20km—if anything effect sizes increase. Columns (3)-(4) restrict our sample to be fully balanced—only 27 plants meet this criteria so our standard errors increase considerably, but coefficients are qualitatively similar to those of the main analysis, albeit statistically insignificant.

1.4.6 Implications of Home Price Analysis

These results provide evidence that households negatively value proximity to natural gas plants and wind turbines and that the average magnitude of this distaste for households nearby the plant (at least as proxied through home values) is roughly similar to the average district-wide gains of allowing these plants to enter into their district. This finding has three important implications that are worth discussing in more depth. First, it suggests that hedonic estimates purporting to capture the disamenity of living near power plants and other large industrial plants must be sure to adjust for jurisdiction by year fixed-effects in order to produce unbiased results. If not properly controlled for,e stimates may be comparing homes nearby the plant that receive both the benefits of the increased tax base and the negative disamenity value from the plant to homes far from the plant that receive less of the tax base benefit (assuming some homes are outside the plant's district) and less of the disamenity. Estimating this joint effect leading to an underestimate of the true disamenity value of these types of projects. We can see that in our setting by estimating the same model as above, but excluding school district by year fixed-effects. Table A.15 shows that when implementing this specification effect sizes decrease by nearly two-thirds as we are now capturing the joint effect of a larger tax base and increased exposure to negative externalities caused by the plant.

An additional related point is that even if hedonic estimates do properly control for district effects (or are estimating values in places where both the control and treatment groups are in the same district), our results suggest that these estimates cannot be used to back-out the change in welfare or even the net change in home prices experienced by a given households. These type of hedonic estimates capture the negative gradient with respect to distance of being near a plant, but our results suggest that these plants also create a positive shift up in home prices for all homes within the plants' district. If the size of the home price increase is larger than the negative price effect caused by a plant's negative externalities, it's possible that a home nearby a plant could experience a net increase in home value, albeit an increase that would be less than that of its neighbor who was further from the plant.

Second, these results have important implications for the distributional consequences of power and other large industrial plant openings. In general, houses within a district receiving a plant, which are far from a plant will benefit from the plant's entrance—these households receive the benefit of increased tax base, but face none of the negative externalities created by the plant. Homes nearby the plant and in the plant's district experience an ambiguous effect on home-values as they both gain the benefits of increased tax base, but also face costs due to their proximity to the plant. Finally, homes nearby, but not in the same district as the plant will experience an unambiguous decrease in home values.

Table A.16 shows the demographics of census tracts within these three categories. All coefficients are relative to being in the plant's district, but more than 15km from the plant. We can see that tracts that are within 5km of the plant, but outside the plant's district—in other words tracts that see an unambiguous loss from the plant—have 7% lower median incomes, 9% lower home values, 5 pp fewer homeowners and 4 pp fewer white residents than tracts that gain an unambiguous benefit (inside the plant's district, but more than 15km away). Homes within the plant's district, but nearby the plant are also much worse off than those faraway, although their median income and median home value is slightly higher than nearby tracts outside the plant's district. These results suggests that despite the fact that plant's benefit host communities by increasing the local tax base, the net effect of these plants is highly unequal. Indeed, because households nearby the plant, and particularly those outside the plant's district are much poorer and more likely to be nonwhite than households inside the plant's district, but far from the plant, these tax base increases may actually serve to exacerbate rather than dampen the inequalities created through these plant's disparate negative impact.

Finally, these results suggest that property tax payments from a plant create a meaningful benefit to local homeowners within the receiving district of a similar magnitude to the negative value created by a plant's pollution and other negative externalities (i.e. noise and shadow flicker for wind turbines). Given the size of this effect, we might expect that access to these revenues could be an important driver in incentivizing local jurisdictions to allow these types of negative-externality producing developments to go forward. In the next section we turn to examining this question explicitly by exploiting cross-sectional and temporal changes in local jurisdictions' ability to access the local tax revenue generated by large industrial plants.

1.5 Effects of Constraining Local Property Tax Revenues on Industrial Development

The previous section showed that property tax payments from large plants are valued by local homeowners. This finding suggests that restricting communities' access to such payments may have large effects on their willingness to allow large negative externality-producing projects to enter into their communities. In this section, we test this question empirically by examining how negative shocks to jurisdictions' ability to use their local tax base to fund local public goods affects their exposure to large manufacturing and power plants.

1.5.1 Data and Empirical Strategy

In this section, we use a series of school finance reforms that occurred across US states between 1970s and 2000s as a plausibly exogenous shock to the marginal value of an additional dollar of local tax base with respect to school spending. As described in Section 2, these reforms were generally aimed at equalizing state education systems and/or increasing the level of education provided by the state's poorest districts. These reforms affected the ability of local school districts to access their local tax base across two dimensions. One, many reforms greatly increased the degree to which state transfers were tied to a district's level of property wealth—following the reform, increases in property wealth would lead to a correspondingly sized decrease in state transfers, which limited (or in the extreme case, eliminated) any benefit districts would receive from an increased tax base. Two, reforms also often instituted tax ceilings and floors, which both limited the amount local districts could tax (and therefore, the amount of revenue they could obtain from their tax base) as well as their ability to cut rates.

Identifying the precise year and type of school finance reforms are difficult (Hoxby, 2001). Many court-decisions led to ostensibly large reforms that in reality had little effect on school finance, while other less-publicized legislative changes led to dramatic shifts in the way in which schools were financed (Lafortune, Rothstein, and Schanzenbach, 2018). In our analysis, we identify reforms by isolating major shifts in the amount of crowd-out or tax limitations embedded within state school funding formulas. In order to find these shocks, we first document state school funding formulas going back to 1962 using *Public School Finance Programs of the United States*, a report published approximately every five years summarizing US states' school finance systems and formulas. Using both the funding formulas and narrative descriptions within these reports, we identified years in which there were large changes in either the crowd-out caused by an additional dollar of tax base or on the level of taxes that a district could charge. Years with substantial shifts in either of these variables were defined as reforms—if a state had multiple major reforms we used the first reform as our event.

Figure A.5 shows summary statistics related to these reforms. The left panel shows the cumulative number of reforms by year and the right panel shows a map of reform states by year of reform. There are two major takeaways. One, the majority of states (34) had major reforms and these states are geographically and demographically diverse. Two, reforms happened nearly continuously between 1970 and 1994—thus results are unlikely to be driven by a specific time-trend like the rapid decline in US manufacturing in the early 2000s.

Using these reform years, we implement a geographically-proximate border pair difference-in-differences analysis, adapting the methodology used by Dube, Lester, and Reich (2010) in their analysis of the employment and wage effects of minimum wage changes. Specifically, we stack all cross-state county pairs whose centroids are within an x mile radius in the United States. In our primary analysis, we use x=60, but show robustness to larger and smaller bandwidths. Because pairs increase exponentially with x, we restrict each county to its twenty closest pairs to ensure results are not driven by more geographically distant counties. Again we show that results are robust to using fewer or all possible pairs.

We define as treated the county in the pair whose state received a reform first (i.e. if one county in the pair had a reform in 1975 and the second county in the pair had a reform in 1990, the 1975 county would be treated). We define the year of the event for the pair as the year of the earliest reform event occurring in either of the pair's states. We drop all years after the second county experiences a reform so as to create an uncontaminated control group. In our primary analysis we maintain a balanced panel by keeping only pairs that have at least 8 years of data before the event and more than 14 years of data after the event, but we show robustness to estimating effects for an unbalanced panel over a broader window. Our specific estimating equation is as follows:

$$Y_{cpst} = \alpha_{cps} + \tau_{pt} + Post_{pst} * Treat_{cst} + \varepsilon_{cpst} \tag{1.5}$$

where Y_{cpst} is the outcome variable in county c, as a part of county pair p, in state s and year t, α_{cps} is a time-invariant fixed-effect for each county in a given county pair, τ_{pt} is a county-pair by year fixed-effect so that all time shocks to common to both counties in a pair are differenced out, $Post_{pst}$ is an indicator for whether or not a reform has occurred, $Treat_{cst}$ is an indicator for whether or not a given county is “treated,” and ε_{cpst} is a mean-zero error term. We implement two-way clustered standard errors at the border pair and state level.

In this analysis, we use large manufacturing establishments, manufacturing employment and power plants as our proxy for local externality producing projects. We view manufacturing plants and establishments as our primary outcome measure for three reasons. First, there are data limitations in the location and timing of power plant openings going back to the early 1960s. We are able to reconstruct the presence of plants that were owned by all utilities still in operation in 1990 and were still in operation or had retired subsequent to 1975. However, data on any plants that may have been owned by utilities that went out of business prior to this period as well as plants or units that retired prior to 1975 are unavailable. Conversely, County Business Patterns provides compehensive data on the stock of manufacturing employment and establishments by county going back to 1964.[34] Second, power plant openings and closings are relatively rare events, while changes in large manufacturing establishments although still rare are far more common. Thus, using manufacturing as an outcome greatly increases our power to detect an effect. Finally, because manufacturing plants bring large regional employment and agglomeration effects, we believe the effect on manufacturing location likely has larger economic implications and is thus of greater interest than power plants. [35]

As described above, our manufacturing data come from County Business Patterns, an annual measure of employment, establishments and payroll within a county by industry type. For this reason, we perform our analysis at the county, not district-level. We have data from 1964 to 2013 for all two-digit NAICS industries.[36] For counties with low levels of employment, employment in certain industries is marked as 0 to maintain privacy. We therefore exclude any

[34]We are in the process of digitizing data to extend our sample back to 1948.

[35]We cannot perform a parallel analysis to Section 3 with manufacturing plants because publicly available data on opening date and location are not available. Further, they would be a less desirable subject for such analysis because their large regional employment and agglomeration effects would make estimating the home price and school finance effects of the tax base changes more challenging. However given the similarity between large manufacturing plants and power plants it is likely that a strong relationship between manufacturing plant openings and the local tax base also exists.

[36]Earlier data is provided for SIC industry codes, which we then convert to NAICS codes.

county-year in which the county has a positive number of establishments, but no employment reported. These are disproportionately small, rural counties and so in our primary analysis we choose to restrict our sample to only counties with greater than 1,000 population in 1970 (prior to any reforms). Results are robust to including all counties in our sample as well as restricting our sample to only counties with greater than 10,000 population in 1970 for which employment is almost universally available for the entire time period of the sample.

We show results from regressions weighted both such that all counties count equally and such that all reforms count equally. Our preferred specification weights all counties equally as this maximizes power, but results are similar (or larger) when weighting all reforms equally. We also show further robustness to not weighting or weighting by population. To account for county duplicates across pairs we cluster all standard errors at the state and border region.

There are two major identifying assumption for this analysis to be interpreted as the causal effect of a shock to reducing local jurisdiction's ability to raise and/or retain property tax revenue. First, we must assume that absent the reform event, manufacturing and power plants in the border counties would evolve on similar paths. This is a priori plausible because these counties are likely exposed to similar geographic shocks. However, we additionally test this assumption in several ways. First, we show estimates in event study form to check for pre-trends. Second, we also show results for manufacturing as a share of all employment and establishments. If effects were driven by broader economic forces then we would not necessarily expect there to be a disproportionate effect among manufacturing establishments, a large producer of negative local externalities. Third, we show that results are robust to a number of different specifications, baseline covariates by year controls and weighting schemes.

A second and more challenging identification assumption is that no other reforms coinciding with the event itself can influence our outcome variables as well as that the reform itself cannot affect our outcome variable through other channels. For instance, many school finance reform events restricted local jurisdictions' ability to raise property tax rates and instead instituted increases in the state sales tax to fund education. If manufacturing firms responded to this changed tax structure, it would change the interpretation of our results. Alternatively, these reforms could have been part of a broader push for progressive legislation that may independently affect locational decisions of plants (i.e. environmental legislation), which would bias our estimats. Finally, these reforms by design increased funding in poor areas—this may have led to changed household location decisions that could influence communities' decision to allow plants to enter through an income effect rather than a price effect.

Although it is difficult to fully rule-out these alternative explanations—and for this reason we see these results as less well-identified than those from the first part of the paper—we attempt to address these concerns in several ways. First, we test if there is within state heterogeneity by local poverty rates—if the observed effects were driven by by the reform increasing funding to poorer communities, we should expect to see larger effects in areas that were low-income prior to the reform. Second, we show suggestive evidence that effects are larger in counties where there is greater overlap between school districts and zoning jurisdictions exactly as theory would predict if effects were driven by the school finance reforms and not othe correlated reforms. Lastly, examining pre-trends can also help identify if effects are being driven by correlated legislation—if so we might expect effects to appear prior to the reforms as not all correlated reforms would be expected to occur in the exact year of the change. In the future, we will perform a set of analyses on single reforms within several states that create large intra-state variation in incentives, which will further help address these concerns.

A final worry may be that our border county design will violate the stable-unit treated value assumption (SUTVA). If counties in reform states are now less receptive to industrial development, prospective plants may be more likely to instead open across the state border in the neighboring county. This is a common feature of all border designs and is unavoidable in our setting. We address this concern in two ways. First, we might expect that the SUTVA concern would be larger when comparing a treated counties to control counties that are very nearby the state border. Accordingly, we estimate if results vary based on a control counties' proximity to the treated state and find no evidence for any such effect.

Second, we can also bound any bias created by this SUTVA violation under the assumption that the reform does not lead to an aggregate increase in the total number of plants across the border pair.[37] If we conservatively assume that every plant that would have opened in the reform county now instead opens in the county across the border, our estimates would be overstated by a factor of 2. Therefore, one-half our observed effect can be thought of as a lower-bound of the true effect.

Finally, it is important to note that although we are looking at changes in school district budgets, it is cities and counties that control local land-use decisions. However, as described in Section 2, in most of the country, there is substantial overlap between school, municipal and county boundaries (Fischel, 2010).[38] Even in places where overlap is incomplete, as long as individuals near the plant within the host municipality and county are in the same school district, we would expect local leaders to internalize these benefits. To the extent that this jurisdictional mismatch leads local municipal and county leaders to discount any school funding benefits, our results would be an underestimate of the effects that would occur if there were shocks to the municipal or county capacity to raise and retain local property tax revenue.

1.5.2 Results

In this section, we examine the effects of shifting local jurisdictions' ability to raise and retain property tax revenue on exposure to externality-producing plants. We examine this question in two ways. We first show that the school finance results obtained in the previous section vary based on their state's level of school finance equalization—this suggests descriptively that these reforms do indeed impact the localized benefit created by plant openings. We then examine in a

[37]Since the reform on net reduces incentives for plant location in the pair, this assumption seems reasonable.

[38]Using 2000 population data, we similarly estimate that in the average county there is an 85% probability that two randomly selected individuals within the same zoning jurisdiction will also be in the same school district.

more causal framework how changing local jurisdictions' ability to raise and retain local property tax revenue affect location patterns of large externality-producing plants using changes induced by school finance reform litigation and legislation.

Heterogeneity in Effects of Plant Openings on School Finance Outcomes

In the previous section, we showed that tax payments produced by these types of projects were economically meaningful and valued by local homeowners. However, many of these openings occurred in states that had already undergone significant equalization reforms, making it likely that these effects are actually much smaller than they would have been in the past, before the reforms were enacted. To test this idea, we examine how the school finance effects of an opening differ by a state's marginal value of tax base (MVTB) with respect to school spending. We proxy for the MVTB in each state by estimating the relationship between a district's total revenue per student and taxable-value per student over time conditional on district and year fixed-effects.[39]

The coefficient on taxable value represents the association between a district's tax-base per student and the total revenue per student available to the district. The higher the value, the more district tax base increases translate into revenue increases on average. Tax limitations and crowd-out will both make this value smaller. Of course, this relationship will also be determined by the extent to which the average district in a state chooses to respond to a tax base shock by shifting property tax rates relative to changing spending. However, given that we saw a much larger spending response than property tax response in Section 3 and because many states have rate floors for eligibility for state funds, we believe this measure is a reasonably good proxy for a district's ability to access its local tax base. We estimate this measure for years 2005 and 2017 to maximize the number of states for which we have data and then apply it to all years in our

[39] Our qualitative method for identifying reforms compares changes within states over time. However comparing the relative stringency of reforms across states is difficult because there is no obvious summary statistic to use to characterize this relationship. In the future, we are working on creating standardized measures of crowd-out and tax limitations for an average district and will use this as an alternate measure of MVTB.

sample.

Figure A.4 shows how the expected tax base impact of a plant differs by its state's MVTB. We would expect this relationship to be positive; if local jurisdictions' can raise and retain a lot of revenue from a plant opening, then the size of a plant's expected fiscal impact should be an important consideration in siting. Conversely, if a jurisdiction cannot retain property tax revenue from a plant then its expected impact is irrelevant. That is precisely the pattern we see here; states with a higher MVTB see location patterns that create higher tax base impacts per student. Although descriptive, this plot provides suggestive evidence that a locality's ability to gain tax benefits from local industry is an important determinant of location choice. This plot also reassuringly shows that our estimated values of MVTB (on the x-axis) are of the magnitude we would expect; in general, states in our sample see a roughly .001 to .01 dollar increase for each dollar of tax base added. Given that the average school district tax rate is roughly 1% and almost all states have some degree of crowd-out, this is precisely the range we would expect.

We next examine how the school finance results from Section 3 vary based on a jurisdiction's estimated MVTB when a plant enters. Table 1.7 shows effects of an opening on total revenue and total expenditures for state-years by above/below median MVTB (Columns 1,2) and estimated MVTB as measured in mills (Columns 3,4). The results are exactly as we would expect. Low equalization (high MVTB) states raise significantly more revenue from plant openings. In low equalization states moving from 25th to 75th percentile in expected tax base impacts (2 log points) leads to a 7% increase in total revenues per student and an 13% increase in total expenditures per student, while in a below median state such an increase leads to no change in revenue and a 2% increase in expenditures. Similarly, a state with an estimated MVTB of 0 sees no change in revenues when moving from a 25th percentile opening to a 75th percentile opening and only a 3% increase in expenditures. Conversely moving to the 75th percentile in MVTB (4.4 mills), leads to a 3.2% increase in revenue and a 8% increase in expenditures. In general, we see that the less able jurisdictions are to raise and retain local property tax revenue the smaller the fiscal benefits of a plant opening.

Effect of School Finance Reforms on Local Industrial Development

In this subsection, we examine in a causal framework how changes in incentives created by school finance reforms impacted local jurisdictions' exposure to large externality-producing plants. We answer this question in the context of school finance reforms and large manufacturing and power plants. We begin by testing whether our qualitatively-identified reforms had meaningful effects on the way schools in a state were funded. Table 1.8 shows these reforms increased the state share of school funding by 9 percentage points, increased state revenue per student by $1,000 and decreased property tax revenue per student by $750, suggesting a massive change in how a state's schools were funded. Further, these reforms were highly progressive; counties with higher poverty rates at baseline saw a much larger increase in state aid. These results provide confirmatory evidence that our qualitatively-identified reforms did indeed lead to a sharp change in incentives for local districts.

We now investigate how this dramatic change in incentives affected exposure to large capital projects.We examine the effect of reform events on both large manufacturing establishments/employment and exposure to large power plants. Table 1.9 shows the main results for the manufacturing analysis. The odd columns show the results using an unbalanced panel of 14 years before and 14 years after the reform and the even columns show the results using a balanced panel of 8 years prior and 14 years after the reform.[40] Because we are measuring the effects of the reform on stocks not flows, we show the effect of being within 5 years of the reform as well as being more than 5 years after a reform as we expect the treatment effect to increase over time. On average, manufacturing employment per capita falls by roughly 6-7 workers per 1,000 population in reform counties or 10%. We also observe a 2 percentage point decline in the share of total employees that work in the manufacturing sector suggesting that this result is not simply driven

[40] Our data begin in 1964 and a large number of reforms occurred in the early 1970s so we cannot include as many pre years in the balanced panel.

by a secular economic decline. Finally, there is a meaningful decline in large manufacturing establishments (as measured by number of employees). Establishments with greater than 250 employees fall by .009 per 1,000 population (20%). Establishments with greater than 500 workers fall by .004 per 1,000 population (20%). All results are robust across samples and are highly statistically significant ($p<.01$).

Table 1.10 show results for the power plant analysis. As described above, we do not have reliable retirement data prior to 1975 so our estimates of generating capacity is somewhat incomplete. With this caveat in mind, the results do provide suggestive evidence of a decline in large plants in reform counties relative to their control neighbors. Specifically, the probability of having a plant of any meaningful size (generating capacity greater than 50MW) falls by 3 percentage points off of a base of 23% and the probability of having a large plant (generating capacity greater than 250MW) falls by 2.5 percentage points off of a base of 15%. Using an inverse-hyperbolic sine transformation we can further see that the total amount of generating capacity within a county falls by 15-20%. All results are marginally statistically significant.

Table A.17 shows the effect of the reform on plant openings and retirements. Openings and retirements are extremely rare events—they occur in less than .5% of county-years and we therefore lack sufficient power to estimate these effects with any kind of precision. Nonetheless, the coefficients all go in the expected directions; reforms lead to a large decrease in openings (in proportional terms) and a somewhat smaller increase in retirements, but given the lack of precision results are generally statistically insignificant.

Together, these results provide evidence that shocks to local jurisdictions' ability to access to their local tax base lead to large changes in siting behavior. However, there are a number of significant identification concerns that may preclude us from interpreting these results causally. In the next subsections, we attempt to test for violations of our identifying assumption across four different domains: omitted variable bias, results driven by other aspects of the reform/correlated reforms, SUTVA violations and weighting/specification/bandwidth choices. While we are unable to fully rule-out many of these violations our results do provide suggestive evidence that it is indeed the change in incentives embedded in the reform that are driving our results.

Omitted Variable Bias

One major identification concern is omitted variable bias; we may be worried that even in the absence of a reform, treated counties would have had different trends in manufacturing employment or power generation than their control neighbors. We begin testing for omitted variable bias by examining whether treatment and control counties have different pre-trends in our outcomes of interest using an event study design. Figure 1.6 shows results dynamically for our manufacturing outcomes, while Figure 1.7 shows dynamic results for our power plant outcomes. In all cases, there are no pre-trends prior to the reform and then a decline following the reform's onset. Because we are examining stocks as an outcome, we would expect the effect to increase in magnitude over time until a new equilibrium is reached; indeed that is precisely the pattern we observe here.

As a second check, we examine whether treatment and control counties differ on baseline characteristics prior to the reform. We are implementing a differences-in-differences analysis and so differences across groups would not be a violation of our identifying assumption per se, but a lack of large differences would still be reassuring that absent a reform these groups of counties would remain on similar trajectories.

Table A.18 test for differences across key demographic and economic characteristics in the Decennial Census preceding the reform. There are no economically or statistically significant differences across any covariates. These results suggest that pre-deterimined differences in baseline characteristics are unlikely to be driving our results. TableA.19 shows differences across outcome variables between treated and control counties in the year of the reform. There are no economically or statistically significant differences in power generation, but treated counties do appear to have a greater number of manufacturing employees and large manufacturing establishments per capita at baseline. Although we did not see any pre-trends in our dynamic analysis, we may still be worried that that this difference in levels between treatment and control counties may be driving our results.

To evaluate this possibility, we perform several tests of the robustness of our results in Tables A.20 (manufacturing employment) and A.21 (large manufacturing establishments). In Columns (1)-(2), we show the results after logging the outcome variable. If results were driven by reforms being correlated with a secular proportional decline in manufacturing, the difference in levels could mechanically create our observed results. Such a scenario is unlikely given that events happened at different times and we did not see any pre-trends in our dynamic analysis, but even more reassuringly we see that results are extremely similar to our main analysis when using a logged dependent variable. Columns (3)-(4) restrict the sample to be only years prior to the start of the massive decline in US manufacturing jobs (in 1997) to ensure that nothing about correlations between this decline and baseline levels of manufacturing are driving our results. Again, results remain largely the same. Columns (5)-(6) restrict our sample to only county pairs that have manufacturing levels per capita within .5 log points in 1964 (the first year of our sample), while Columns (7)-(8) use a .25 log point cut-off. Despite significantly reducing our sample size, results remain qualitatively similar. Finally, Columns (9)-(10) exclude border regions whose pre-reform differences in manufacturing employment are large positive outliers. In this sample, there are no economically or statistically significant baseline differences in manufacturing employment and yet results persist suggesting again that these baseline differences are very unlikely to be driving our main results.

Finally, if results were driven by other trends we might also expect changes to occur in other areas of the economy aside from manufacturing. We expect the effects of these reforms to be largest among large polluting plants for two reasons. One, these plants are often the most valuable pieces of non-residential real estate in a given district. While properties like warehouses or office buildings also create tax revenue because they are typically less capital-intensive than large plant, each individual property contributes much less to the local tax base and therefore the tax benefits are less likely to be an important reason for approval. Two, these plants create large local externalities that are likely to induce substantial local opposition to the project in the absence of compensation.

We test this supposition in several ways. Table A.22 shows the effects of the reforms on non-manufacturing employment and industries, while Table A.23 breaks down the effects of the reform by industry type. There are no significant effects of the reform across non-manufacturing industries providing further evidence that results are not driven by a secular economic decline in treated counties. Further, there is suggestive evidence for employment declines that in other sectors that may have large externality producing projects such as the mining/extraction or transportation/utility sectors although these effects are not statistically significant. Together, these results suggest that differential trends among treated and control counties are not likely to be driving our results.

Simultaneous Shocks

A second and harder to rule-out identification concern is that results are driven by shocks that occurred simultaneous to the change in incentives. We can divide this concern into two parts: other aspects of the school finance reform are driving the results and/or reforms correlated with the adoption of school finance are causing the results. We address each of these possibilities in turn.

School finance reforms typically increased spending in the poorest districts and weakly decreased spending in the richest districts. Such changes in spending could induce sorting, increase land-values in poor districts or the marginal value of additional tax base because of diminishing marginal returns all of which would lead polluting industries to decline in low-income areas of reform states for reasons other then the change in incentives. However, if these dynamics were driving our observed results, we would expect effects of the reforms to be much larger in high-poverty relative to low-poverty districts. We test this hypothesis empirically in Table 1.11. This table shows that if anything, effects are larger in low-poverty relative to high-poverty counties suggesting that these other aspects of the reforms are unlikely to be driving the observed results.

An additional concern is that these reforms may have occurred as part of a suite of progressive legislation that may independently have had effects on the location of polluting industries (i.e. laws requiring strict environmental or community impact assessments prior to development). We think this is unlikely to be the cause of the observed results for several reasons. First, many school finance reforms were the results of lawsuits alleging that existing funding structures violated the state constitution (Jackson, Johnson, and Persico, 2014) and not part of a broader legislative push. Second, reforms happened across a diverse range of states and over a long time period. Third, we do not observe major changes outside of the manufacturing and power generation industries—thus, such reforms would have to be highly targeted to achieve such an effect. The most likely type of reforms that could lead to this pattern are environmental reforms, but in the vast majority of states most major environmental initiatives have been federal. However, we are in the process of collecting information on major environmental reforms across US states to empirically assess this possibility.

In Table A.24 we provide a further empirical check of this assumption. As discussed above, the studied reforms affected school districts, but the jurisdictions typically making zoning decisions are cities and counties. Thus, we might expect that the more closely school districts align with zoning jurisdictions, the stronger the effect a reform should have. School districts and zoning jurisdictions can be mismatched on two dimensions; a single zoning jurisdiction can be spread over multiple school districts and similarly a single school district can be spread over multiple zoning jurisdictions. Both dimensions should matter for the effect of the reform. If a zoning jurisdiction is spread over multiple school districts fewer residents will benefit from a given plant opening. If a school district is spread over multiple zoning jurisdictions, the amount of benefit will be diluted within each zoning jurisdiction. Thus, our variable of interest is an interaction term between the average probability that two given residents of a county's zoning jurisdiction will in the same school district and the average probability that two given residents of a school district will be in the same zoning jurisdiction.

Table A.24 shows some suggestive evidence that effects are larger in areas with greater overlap between school districts and zoning jurisdictions, although effects are only significant in the case of employment. The results in this table include state by year fixed-effects and so are identified wholly off of differences in effect-sizes within treated counties in the same reform state. Because no other major state environmental laws is likely to target counties in states with disproportionate overlap between school and zoning jurisdictions, we believe this pattern provides further suggestive evidence for the change in incentives as the main mechanism driving these results.

Stable Unit Treatment Value Assumption Violation

An additional concern may be that our analysis may violate the stable unit treatment value assumption. The idea here would be that if a plant's previously optimal location was in a given county, its next most optimal county after the reform may be directly across state lines in the control county. We address this concern in two ways. First, we can attempt to put a lower bound on the treatment effect if SUTVA is operative. We first assume that the treatment does not increase the total number of plants locating in a given pair; this assumption seems plausible as the treatment is making it less likely on average that the pair will want a plant. Under this assumption, a violation of SUTVA could overstate our result by at most a factor of 2 if all plants from the treated county moved to the control county. All results would remain highly economically significant under this assumption.

In addition to bounding, we attempt to test this supposition empirically in Table A.25. Specifically, we interact our difference-in-differences estimator with a variable equal to the distance of the control county to the treated state border (odd columns) or exclude control counties that are within 45 miles of the border (even columns). We expand our sample to include all county pairs that are within 90 miles of each other to provide sufficient power to test these hypotheses. We find no evidence that effects differ by distance to border or are significantly smaller after excluding counties nearby the border. These results provide some supportive evidence that SUTVA violations are not a major determinant of our observed results.

Weighting/Specification/Bandwidth

Finally, we perform several checks to ensure that our weighting, sample selection, or other specification choices are not driving our results. Table A.26 and A.27 show results when weighting by reform event or population. We see that results for both manufacturing and power generation are if anything larger when weighting by population. When weighting by reform state results are broadly similar although standard errors increase as the effects of several states are estimated imprecisely. Table A.28, A.30 and A.29 show results using different distance bandwidths. Results are broadly similar regardless of whether we restrict the bandwidth to 30 miles or increase it to 90 miles.

We might also be concerned that something in our reform identification procedure itself is driving results. To address this, we perform the same analysis using the first school finance reform identified for each state in Jackson, Johnson, and Persico (2014) analysis of school finance reform's effects on low-income educational achievement and labor market outcomes. We would expect the effects using these reforms to be smaller because not all reforms used by Jackson, Johnson, and Persico (2014) change property tax incentives—some simply equalize spending through lump-sum transfers or other techniques, which do not change local control over property taxation. The observed results are consistent with this pattern; using the Jackson, Johnson, and Persico (2014) reforms we see directionally similar results that are also statistically significant, but the magnitudes are often smaller than those found when using our qualitatively identified reforms. These results provide additional reassurance that nothing particular to our qualitative reform-identification process is causing our results.

Finally, we can identify the effects separately for each reform state to ensure that no one state is behind our results. Figure A.7 shows the distribution of effects across states. This figure shows that almost all states have negative treatment effects suggesting that results are not driven by one or two states. Treatment effect sizes are not correlated with reform timing nor geographic region of the country. In sum, these results suggest that the observed effects are unlikely to be driven by specification, bandwidth, weighting or other similar factors.

1.6 Conclusion

Large capital projects create substantial external benefits and costs. When these costs and benefits occur on different spatial scales in the presence of local control over land-use, inefficiencies can emerge. In this paper, we study how access to local property taxation may change local jurisdictions' willingness to allow externality-producing projects enter their communities. We first show that in addition to the negative externalities imposed by plants, nearby residents also have the potential to experience significant gains from plant openings in the form of increased property tax payments. The average opening leads to a 10% increase in the tax base on average. This increased tax base further caused increased educational spending, used largely on capital expenditures. There is also a small decrease in local property tax rates.

We next show that local homeowners value this increased educational spending. After the plant opens, homes within the receiving district increase by 3%-7% in value for an average opening relative to similar homes just across the border. This increase is of a similar magnitude to the decrease in home prices caused by the plant for nearby residents suggesting that property taxation of large plants has important distributional consequences for who is helped and who is harmed by their construction.

We finally examine how changing local jurisdictions' ability to access these tax revenue benefits affects their openness to externality-producing projects. To investigate this question empirically, we use plausibly exogenous changes in crowd-out and tax limitations caused by school finance litigation and legislation. We show that following a reform, manufacturing employment and large establishments fall by 10-15% suggesting that the benefits provided by property tax payments from these entities can be an important driver of local industrial development.

These results also suggest that reforms that restrict a local government's ability to raise revenue from their tax base may have significant unintended consequence for local land-use. This is a feature of many common state-level policies including school finance reforms, municipal and county revenue sharing systems and property tax limitations. However, the welfare implications of this shift are not clear. Depending on the relative distribution of local and social costs and benefits, limiting the property taxation benefits of these projects may either increase or decrease efficiency. Better understanding this trade-off is essential when considering the design and reform of state-level programs that infringe upon local property taxation. In future work, we will strive to both characterize the efficiency loss and gains created by this policy as well as investigate how changes to boundaries of taxing jurisdictions can affect this trade-off.

Chapter 1, in full, is currently being prepared for submission for publication of the material. Fraenkel, Rebecca; Krumholz;Samuel. “Property Taxation as Compensation for Local Exeternalities: Evidence from Large Plants”. The dissertation author was a primary investigator and author of this material.

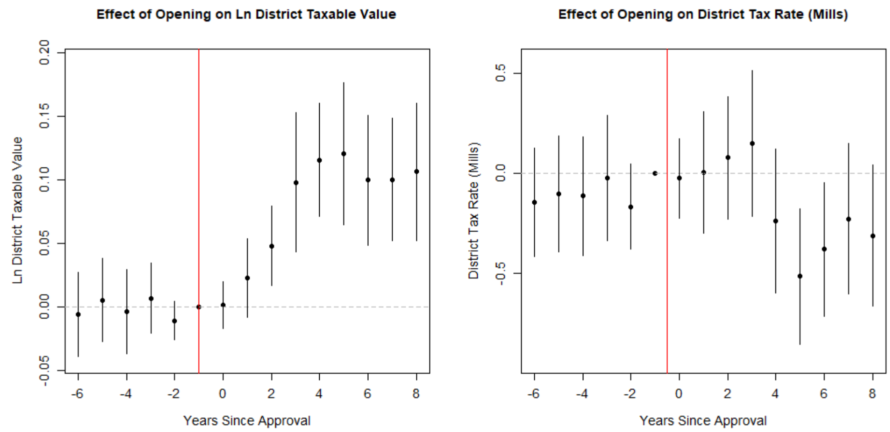


Figure 1.1: Effect of Opening on Taxable Value Per Student and Property Tax Rates

This figure shows the effect of a plant opening on the natural log of district taxable value per student (left) or property tax rates in mills (right). Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between indicators for years since approval (-1 is the omitted category) and whether or not a district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. Only openings with greater than $10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. District taxable values and property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Plant opening data comes from the Energy Information Administration (EIA).

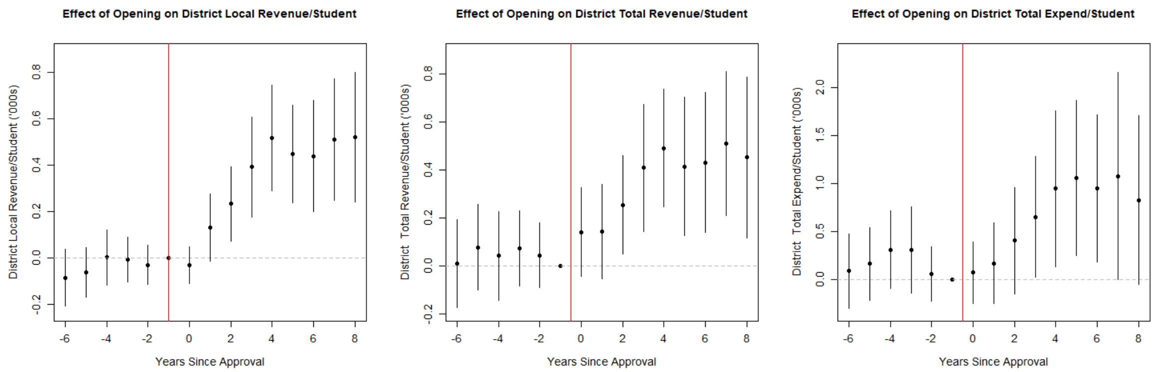


Figure 1.2: Effect of Opening on Local Revenues/Student, Total Revenues/Student and Total Expenditure/Student

This figure shows the effect of a plant opening on district revenues and expenditures per student. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between indicators for years since approval (-1 is the omitted category) and whether or not a district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. Revenue and expenditure data came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

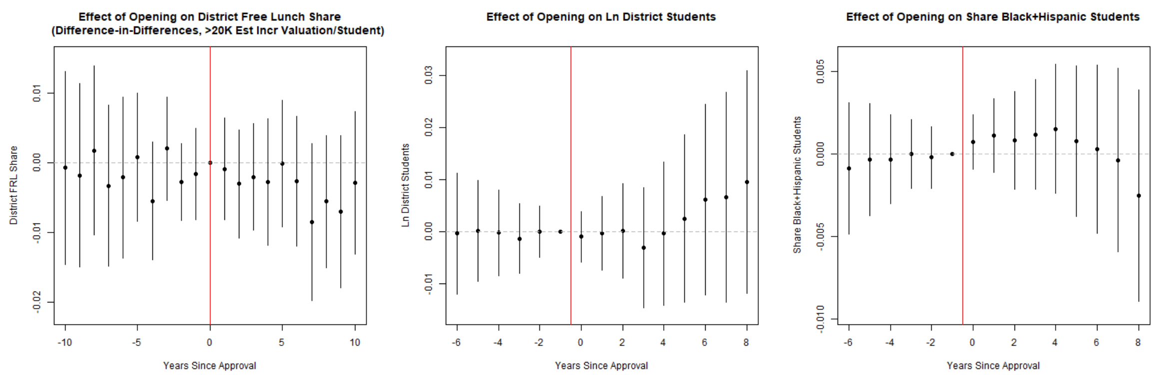


Figure 1.3: Differences in Key Demographic Groups Before and After Openings

This figure shows the effect of a plant opening on share of students with free and reduced lunch (FRL, 1998-2018 only), log enrollment and share of black and hispanic students. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between indicators for years since approval (-1 is the omitted category) nnd whether or not a district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. Demographic data came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

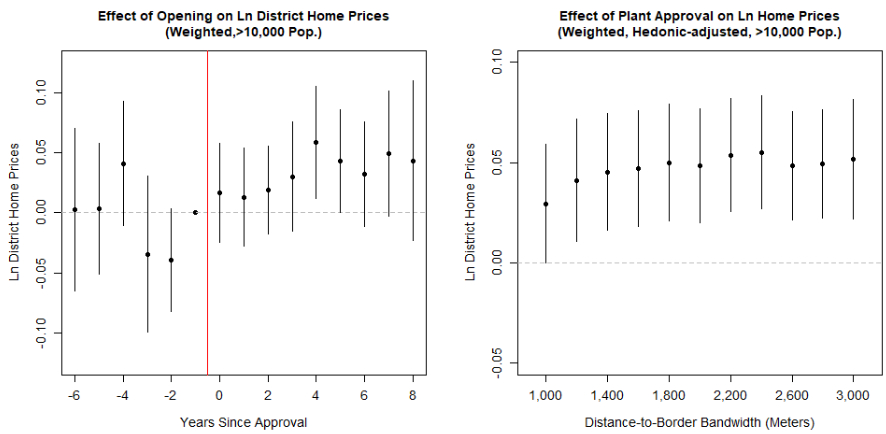


Figure 1.4: Effect of Opening on Host-District Home Prices

This figure shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. In the left panel, we only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for years since approval. In right panel, we show the coefficient of a regression of ln prices on an indicator for whether a house was in a plant receiving district with an indicator for years 3-8 since plant approval using different border bandwidths. In both regressions additional controls include border pair by year by month fixed-effects and .004 degree latitude x .004 degree longitude x year fixed-effects. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below \$5,000 or greater than \$1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than \$10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

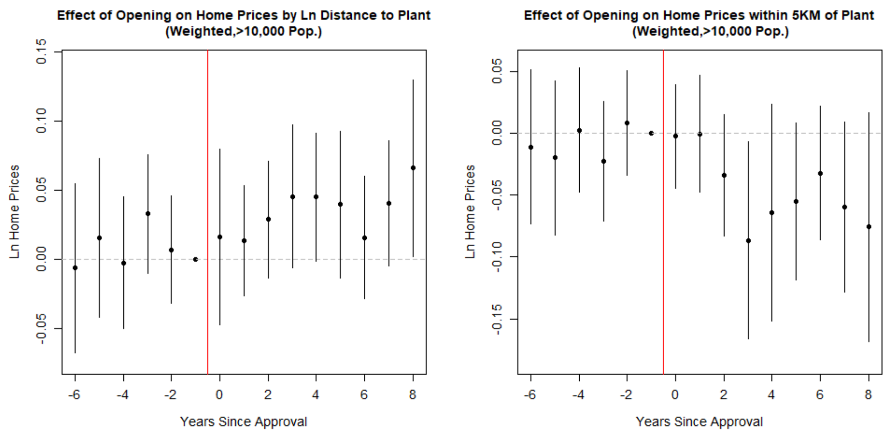


Figure 1.5: Effect of Opening on Nearby Home Prices

This figure shows the effect of a plant opening on log local housing prices for nearby homes. All regressions include controls for plant by year by month fixed-effects and .004 degree latitude by .004 degree longitude cell fixed effects. Coefficients are the interaction between an indicator for years since plant approval (-1 is the omitted variable) and two distance metrics: ln distance from plant (on left) and an indicator for being less than 5km from the plant (on right). Homes more than 20km away and closer than .5km from the plant are dropped and the regression on the right also includes indicators for being 5km-10km from the plant x years since approval. Plants in districts with fewer than 10,000 population are excluded as there are too few annual transactions to create consistent home price estimates. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales attached to a given plant in each year. Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

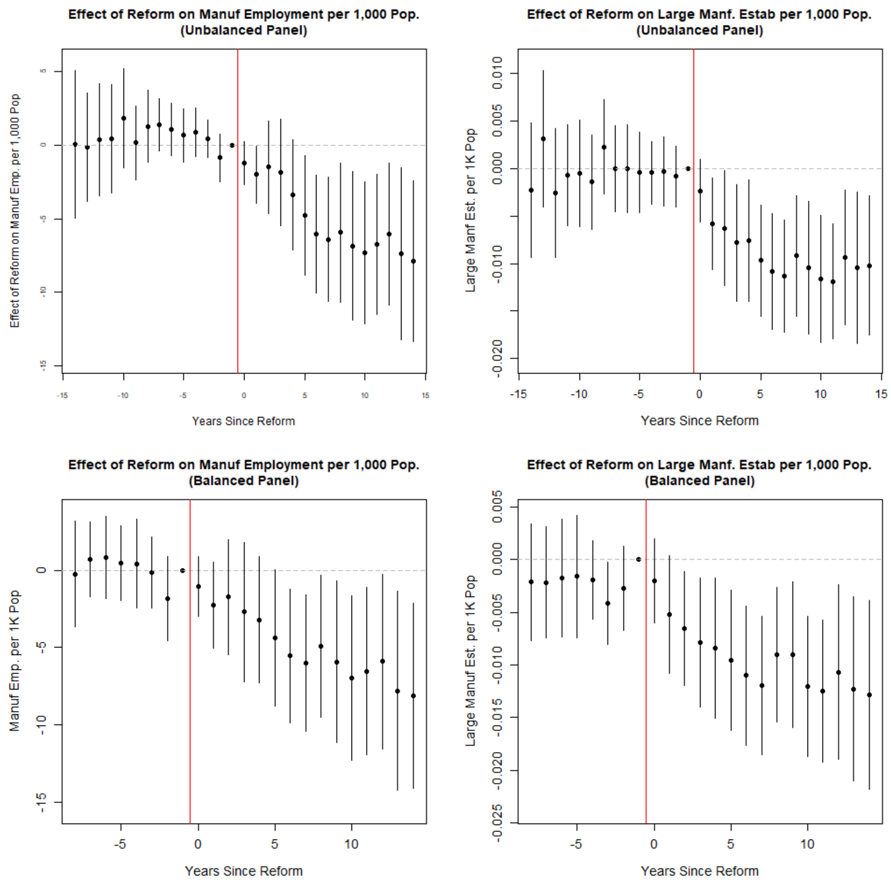


Figure 1.6: Effect of School Finance Reform on Large Manufacturing Establishments and Manufacturing Employment

This figure shows the effect of a school finance reform on manufacturing outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. In the case where both members of a pair were treated, all years after the second event occurred were dropped. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (the year prior to reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing employment data come from County Business Patterns. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. We exclude outcome values greater than the 99th percentile as outliers. Standard errors are clustered at the state and state border pair levels.

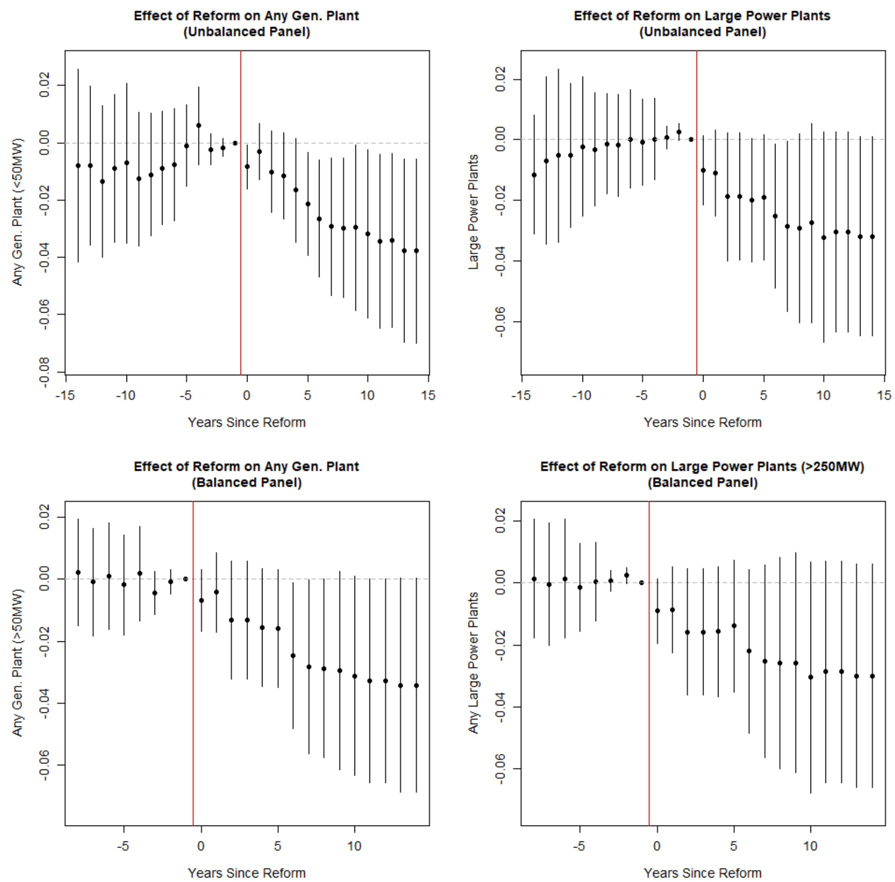


Figure 1.7: Effect of School Finance Reform on Power Plant Openings

This figure shows the effect of a school finance reform on power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. In the case where both members of a pair were treated, all years after the second event occurred were dropped. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (the year prior to reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing employment data come from County Business Patterns. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. We exclude outcome values greater than the 99th percentile as outliers. Standard errors are clustered at the state and state border pair levels.

Table 1.1: Effects of Plant Opening on District Tax Base and School Finance Outcomes

VARIABLES	(1) Tax Value	(2) Rate	(3) Loc Rev	(4) Tot Rev	(5) Tot Exp
Treat x Post Yrs 0-2	0.0260**	0.110	0.141**	0.138*	0.0602
	(0.0130)	(0.113)	(0.0599)	(0.0792)	(0.175)
Treat x Post Yrs 3-8	0.108***	-0.167	0.501***	0.409***	0.765**
	(0.0236)	(0.134)	(0.111)	(0.114)	(0.344)
Observations	21,240	21,092	38,984	38,962	39,882
R^2	0.966	0.973	0.955	0.937	0.853
Pair x Year FE	Y	Y	Y	Y	Y
Sample	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small
Dep. Var. Mean	580887	10.56	4.991	11.79	12.15

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on a district's taxable value per student, tax rate (mills), local revenue per student ('000s/student), total revenue per student ('000s/student) and total expenditures per student ('000s/student). Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. . Only openings with greater than \$10,000 in expected tax base per student are included (≈ 10% of openings are dropped). We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. District property tax rates and tax bases were hand-collected from state Department of Education and Department of Revenue's annual reports. Revenue data came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table 1.2: Effects of Plant Opening on School Finance Outcomes by Revenue Source

VARIABLES	(1) Ptax /Stud	(2) PrntGov /Stud	(3) UnspecLoc /Stud	(4) OthLoc Stud	(5) StFormAid /Stud	(6) OthSt /Stud	(7) Fed /Stud
Treat x Post Yr 0-2	0.0697*	0.0315*	0.0287	0.0108	-0.0636	0.00417	-0.00616
	(0.0418)	(0.0175)	(0.0191)	(0.0175)	(0.0421)	(0.0349)	(0.0127)
Treat x Post Yr 3-8	0.338***	0.0666**	0.0714***	0.0270	-0.134***	0.0110	-0.0224
	(0.0913)	(0.0269)	(0.0249)	(0.0267)	(0.0504)	(0.0282)	(0.0157)
Observations	38,818	38,818	38,818	38,818	39,262	39,262	39,056
R^2	0.961	0.983	0.677	0.944	0.964	0.906	0.937
Pair x Year FE	Y	Y	Y	Y	Y	Y	Y
Sample	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small
Dep. Var. Mean	3.473	0.565	0.210	1.201	4.294	1.537	0.874

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on a district's revenue streams measured on a per-student basis. Uncategorized local and state payments represent payments that do not fit into the NCES categories and often encompass payments in leiu of taxation (PILOT). Other local payments are largely made up of sales and income taxes. Other state revenues are transfers for state-mandated programs like transportation, special education or English language learners. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. Only openings with greater than $10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. Revenue data came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table 1.3: Effects of Plant Opening on Debt and Expenditures by Type

VARIABLES	(1) LTD /Stud	(2) Instr Sal. /Stud	(3) Cap /Stud	(4) Oth /Stud
Treat x Post Yrs 0-2	0.248	-0.0134	0.0920	-0.0184
	(0.279)	(0.0258)	(0.129)	(0.110)
Treat x Post Yrs 3-8	0.863*	0.0430	0.402***	0.319
	(0.462)	(0.0381)	(0.142)	(0.247)
Observations	39,712	39,882	39,882	39,882
R^2	0.829	0.974	0.597	0.878
Pair x Year FE	Y	Y	Y	Y
Dep. Var. Mean	5.819	6.103	1.351	4.796

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on a district's outstanding long-term debt per student, total instructional personnel salaries per student, total capital expenditures per student and all other expenditures per student. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. Only openings with greater than $10,000 in expected tax base per student are included (≈ 10% of openings are dropped). We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. Schol finance data came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table 1.4: Effects of Plant Opening on Home Prices

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price
Treat x Post Yrs 0-2	0.0220	0.0370	0.00696	0.0290*	0.0552**	-0.00175
	(0.0145)	(0.0227)	(0.0166)	(0.0152)	(0.0247)	(0.0152)
Treat x Post Yrs 3-8	0.0485***	0.0700***	0.0275	0.0508***	0.0718**	0.0285
	(0.0143)	(0.0207)	(0.0185)	(0.0175)	(0.0291)	(0.0168)
Observations	501,699	198,576	303,123	538,353	223,992	314,361
R^2	0.708	0.639	0.804	0.727	0.685	0.802
Weighted	Y	Y	Y	Y	Y	Y
Spec	.4km grid	.4Km Grid	.4km Grid	.4km grid	.4km grid	.4km grid
Sample Pop Excl	>10K Pop	>10k Pop	>10k Pop	All	All	All
Sample Geo Excl	All	No CA	CA-Only	All	No-CA	CA-only

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and .004 degree latitude x .004 degree longitude x year fixed-effects. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. A population exclusion of greater than 10,000 population means that only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than $10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table 1.5: Effects of Plant Opening on Home Prices: Different Expected Tax Base Per Student Cut-Offs

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price
Treat x Post Yrs 0-2	0.00945	0.0403**	0.0611**	-0.00189	0.0208*	0.0364***
	(0.00929)	(0.0183)	(0.0289)	(0.00752)	(0.0104)	(0.00943)
Treat x Post Yrs 3-8	0.0309***	0.0708***	0.105**	0.00650	0.0498**	0.0687***
	(0.0104)	(0.0241)	(0.0430)	(0.00967)	(0.0207)	(0.0171)
Observations	1,107,048	143,277	61,487	1,156,284	165,509	75,216
R^2	0.715	0.701	0.704	0.716	0.639	0.640
Weighted	Y	Y	Y	N	N	N
Sample	>10k Pop	>10k Pop	>10k Pop	>10k Pop	>10k Pop	>10k Pop
Exp Tax Base Cut-off	All	75K/Stud	150K/Stud	All	75K/Stud	150K/Stud
Avg Exp Tax Base/Stud	1.076	2.685	4.130	0.344	1.871	2.878

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and .004 degree latitude x .004 degree longitude x year fixed-effects. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation in Columns (1)-(3) are weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table 1.6: Effects of Plant Opening on Nearby Home Prices: Spatial Difference-in-Differences

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price
5-10kM Away x Post Yrs 0-2	-0.00546	-0.00377				
	(0.0118)	(0.0108)				
5-10 Km Away x Post Yrs 3-8	-0.0288*	-0.0185				
	(0.0168)	(0.0119)				
<5km away x Post Yrs 0-2	-0.00633	-0.01277				
	(0.0130)	(0.0108)				
<5km away x Post Yrs 3-8	-0.0579**	-0.0329*				
	(0.0252)	(0.0168)				
Ln Dist x Post Yrs 0-2			0.00864	0.00863		
			(0.0126)	(0.00980)		
Ln Dist x Post Yrs 3-8			0.0313*	0.0240**		
			(0.0157)	(0.0115)		
5-10Km Away x Post Yrs 0-2 x Nameplate ('00 MW)					-0.00284	-0.00393**
					(0.00219)	(0.00173)
5-10Km Away x Post Yrs 0-2 x Nameplate ('00 MW)					-0.00895***	-0.00878***
					(0.00309)	(0.00123)
< 5Km Away x Post Yrs 0-2 x Nameplate ('00 MW)					-0.00620***	-0.00520***
					(0.00226)	(0.00117)
< 5Km Away x Post Yrs 3-8 x Nameplate ('00 MW)					-0.0150***	-0.00920***
					(0.00486)	(0.00190)
Observations	2,228,378	2,352,691	2,228,378	2,352,691	3,964,489	4,107,232
R^2	0.683	0.705	0.683	0.701	0.690	0.698
Size Cutoff	>100MW	>100MW	>100MW	>100MW	N	N
Hedonics	Y	Y	Y	Y	Y	Y
Max Dist	20km	20km	20km	20km	20km	20km
Weighted	Y	N	Y	N	Y	N

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on local housing prices for nearby homes. All regressions include plant district by year, .004 degree longitude x .004 degree latitude bins, and district x year fixed-effects.The outcome variable is residualized for hedonic by state controls. These include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms, square footage (500 sq ft bins), and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Only homes within 20km of the opening plant are included. Only plants who are in districts with a population greater than 10,000 are included in weighted analysis. Standard errors are clustered at the plant district level. Only openings with at least two years of pre and two years of post data are included. All housing data come from the Zillow ZTRAX database—sales that are less than $5,000 or greater than $1,500,000 are excluded as outliers. Weighted specifications are weighted by the inverse of the number of sales within 20km of the plants in a given year.

Table 1.7: Differential Effects of Plant Opening on Key School Finance Variables by State Equalization Status

VARIABLES	(1) TotRev	(2) TotRev	(3) TotExp	(4) TotExp
Treat x Post Yrs 0-2 x Ln Exp Tax Base Incr/Stud	-0.00229	-0.00256	-0.00900	-0.00811
	(0.00458)	(0.00487)	(0.00672)	(0.00694)
Treat x Post Yrs 3-8 x Ln Exp Tax Base Incr/Stud	-0.00173	0.00367	0.0108*	0.0147**
	(0.00431)	(0.00460)	(0.00609)	(0.00678)
Treat x Post Yrs 0-2 x Ln Exp Tax Base Incr/Stud x > Med MVTB	0.0234***		0.0351***	
	(0.00682)		(0.0109)	
Treat x Post Yrs 3-8 x Ln Exp Tax Base Incr/Stud x > Med MVTB	0.0384***		0.0543***	
	(0.0110)		(0.0158)	
Treat x Post Yrs 0-2 x Ln Exp Tax Base Incr/Stud x MVTB (Mills)		0.00305***		0.00447***
		(0.000966)		(0.00154)
Treat x Post Yrs 3-8 x Ln Exp Tax Base Incr/Stud x MVTB (Mills)		0.00399***		0.00650***
		(0.00150)		(0.00219)
Observations	25,106	25,106	25,708	25,708
R^2	0.939	0.938	0.886	0.886
Sample	All	All	All	All
Dep. Var. Mean	11.58	11.65	11.89	11.89

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on various district-level school finance outcomes. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between indicators for whether a year is after an approval and whether or not a district receives a plant, log expected tax base per student and various quantiles of estimated size of the increase in tax base per student the plant will provide. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. We proxy for district's estimated marginal value of tax base (MVTB) using a coefficient derived from a state-specific regression of state and local revenue per student on taxable value per student with district and year fixed-effects. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table 1.8: School Finance Reform and County School Revenue by Source: County Pairs Design

VARIABLES	(1) StShare	(2) StRev /Stud	(3) StRev /Stud	(4) PtaxRev /Stud	(5) PtaxRev /Stud
Treat x Post Yr $\leq$5	0.0605***	0.794***		-0.373	
	(0.0176)	(0.199)		(0.232)	
Treat x Post Yr >5	0.0849***	1.079***		-0.701*	
	(0.0213)	(0.247)		(0.357)	
Treat x Post Yr $\leq$5 x BL Poverty Rate			3.344***		-0.394
			(1.023)		(0.844)
Treat x Post Yr >5 x BL Poverty Rate			3.510***		-1.091
			(0.953)		(1.087)
Observations	40,296	40,296	40,102	40,580	40,388
R^2	0.949	0.924	0.969	0.914	0.970
Unbalanced	Y	Y	Y	Y	Y
Dep. Var. Mean	0.369	3.507	3.507	2.688	2.688

Two-way standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on school finance outcomes using a county border pair difference-in-differences design. All revenue outcomes are measured in ’000s of dollars. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Baseline poverty rate is the 1970 county poverty rate. Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. All school finance data come from the Census of Governments (COG) and National Center for Economic Statistics (NCES). All statistics are aggregated to the county level, with a district being assigned to its primary county as defined by COG/NCES. Only districts with data missing in fewer than 10% of years are included to insure constancy of the sample within each county over time. Counties with fewer than 1,000 population in 1970 are excluded. Our sample consists of 14 years before and after the reform year.

Table 1.9: School Finance Reform and Manufacturing Presence: County Pairs Design

VARIABLES	(1) EmpManf /1K Pop	(2) EmpManf /1K Pop	(3) EmpManf /TtlEmp	(4) EmpManf /TtlEmp	(5) >250 Emp Est /1K Pop	(6) >250 Emp Est /1K Pop	(7) >500 Emp Est /1K Pop	(8) >500 Emp Est /1K Pop
Treat x Post $\leq$5 Years	-3.159**	-2.529	-0.00996*	-0.0101*	-0.00673**	-0.00520*	-0.000869	0.000595
	(1.500)	(1.595)	(0.00514)	(0.00501)	(0.00265)	(0.00258)	(0.00122)	(0.00123)
Treat x Post >5 Years	-7.315***	-6.434***	-0.0206***	-0.0185***	-0.0108***	-0.00938***	-0.00442**	-0.00333*
	(2.096)	(2.263)	(0.00643)	(0.00641)	(0.00314)	(0.00317)	(0.00167)	(0.00175)
Observations	164,626	91,218	165,546	91,644	206,702	119,582	205,986	119,586
R^2	0.955	0.959	0.957	0.964	0.902	0.913	0.865	0.874
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	74.09	73.16	0.324	0.301	0.0649	0.0619	0.0206	0.0237

Two-way clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing employment data come from County Business Patterns. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. We exclude outcome values greater than the 99th percentile as outliers.

Table 1.10: School Finance Reform and Power Plants: County Pairs Design

VARIABLES	(1) Any >50MW	(2) Any >50MW	(3) Ttl >50MW	(4) Ttl >50MW	(5) Any >250MW	(6) Any >250MW	(7) IHS(MW)	(8) IHS(MW)
Treat x Post Yr $\leq$5	-0.00583	-0.0110	-0.00306	-0.0167	-0.00370	-0.0137	-0.0207	-0.0981
	(0.00892)	(0.00979)	(0.0102)	(0.0126)	(0.00645)	(0.00854)	(0.0612)	(0.0709)
Post x Treat Yr >5	-0.0281*	-0.0302*	-0.0276	-0.0379	-0.0230*	-0.0279*	-0.172*	-0.214*
	(0.0141)	(0.0164)	(0.0218)	(0.0252)	(0.0134)	(0.0158)	(0.0957)	(0.112)
Observations	212,448	122,938	212,448	122,938	212,448	122,938	212,448	122,938
R^2	0.965	0.969	0.975	0.977	0.960	0.963	0.965	0.969
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Unbalanced	Unbalanced	Unbalanced	Unbalanced	Unbalanced
Dep. Var Mean	0.226	0.223	0.326	0.356	0.145	0.150	215.4	229.5

Two-way clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Power plant data come from EIA Form 860. For plants that retired prior to 1990, only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event.

Table 1.11: School Finance Reform, Manufacturing Presence and Power Plants by Baseline Poverty

VARIABLES	(1) EmpManf	(2) EmpManf	(3) LrgManfEst	(4) LrgManfEst	(5) AnyGen(50MW)	(6) AnyGen(50MW)	(7) AnyGen(250MW)	(8) AnyGen(250MW)
Pov Status	> Med	< Med	> Med	< Med	> Med	< Med	> Med	< Med
Treat x Post Yr $\leq$5	-1.527	-4.215*	-0.00560	-0.00525*	-0.0110	-0.0111	-0.00578	-0.0218
	(2.598)	(2.151)	(0.00403)	(0.00304)	(0.00981)	(0.0171)	(0.00712)	(0.0147)
Treat x Post Yr >5	-4.103	-8.350**	-0.00984**	-0.00793**	-0.0166	-0.0443	-0.00760	-0.0489
	(3.729)	(3.277)	(0.00444)	(0.00361)	(0.0108)	(0.0311)	(0.00748)	(0.0293)
Observations	46,760	47,278	63,706	58,170	64,506	58,432	64,506	58,432
R^2	0.956	0.965	0.913	0.927	0.982	0.958	0.979	0.950
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	0.226	0.223	0.326	0.356	0.145	0.150	215.4	229.5

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. For plants retiring prior to 1990, only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

A.1 Tables and Figures

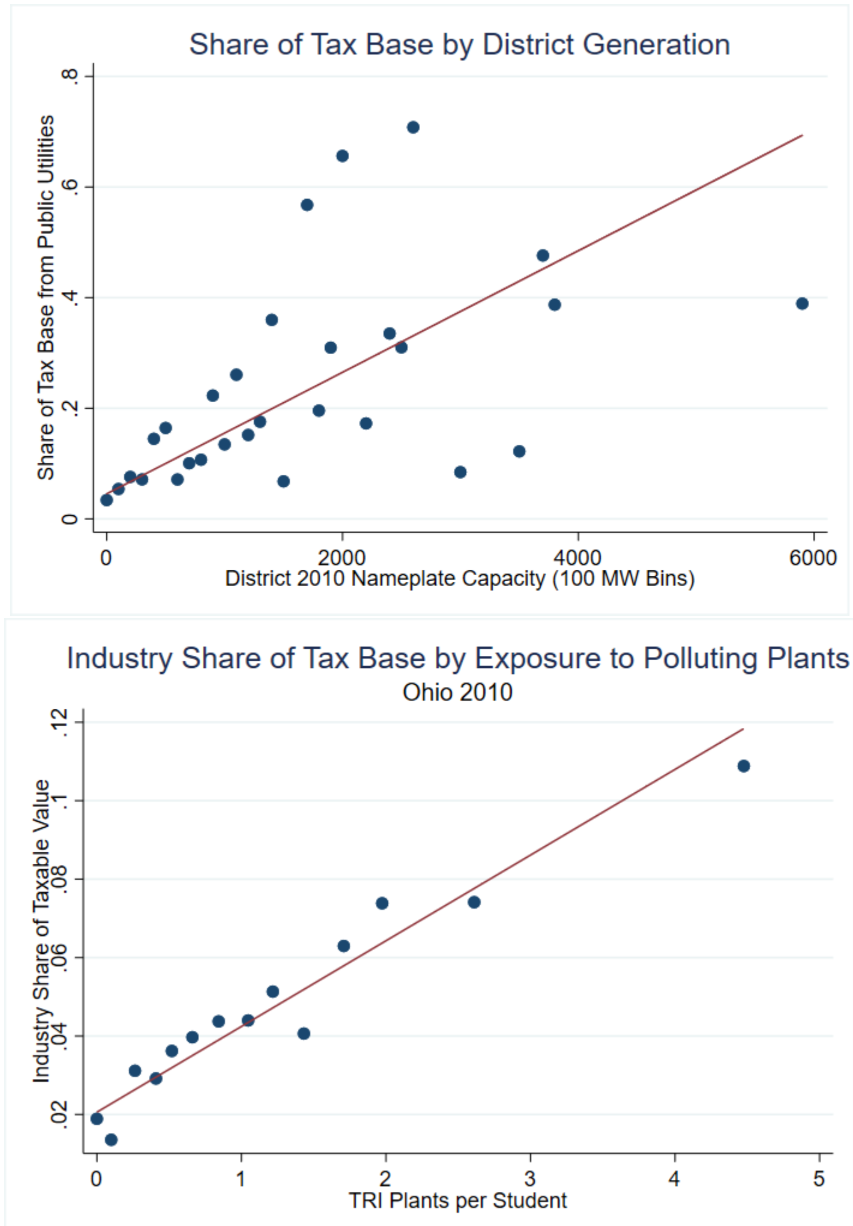


Figure A.1: Utility Share of School District Tax Base by District Generation Level

The top panel of this figure shows the proportion of a district's 2010 tax base that is made up of utility property as a function of the amount of generating capacity located in a district. All data are from 2010 and come from the 8 states with utility valuation data available: Connecticut, Georgia, Iowa, Minnesota, Ohio, Oklahoma, Oregon, and Washington. Hydroelectric generation is excluded as most dams are federally-owned and pay payments-in-lieu-of-taxes (PILOT) rather than property taxes. The bottom panel of this figure shows the proportion of a district's tax base that is made up by industry as a function of Toxic Release Inventory (TRI) plants per student within a district . Data is for Ohio only and from 2010.

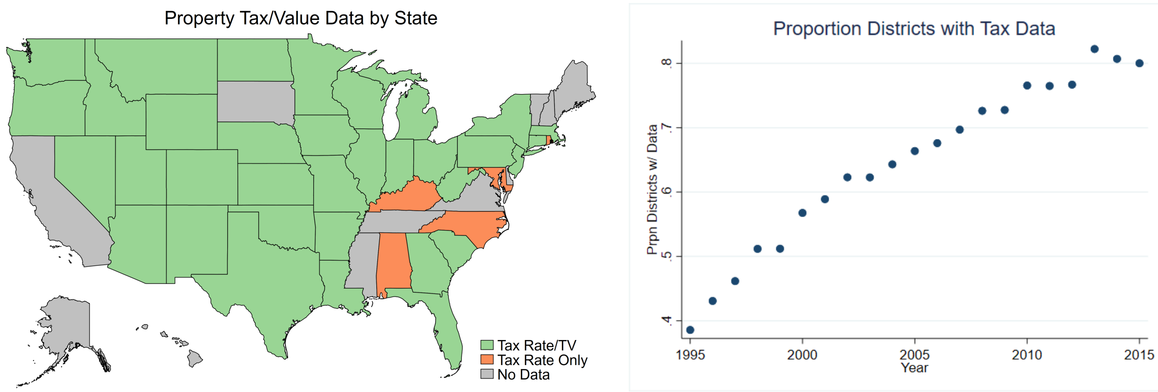


Figure A.2: Data Coverage of Property Tax Rates and Taxable Value

This figure shows coverage of district-level data on school district property tax rates and total taxable value. The figure on the left shows geographic coverage—"Tax Rate/TV" denotes that a state has both tax rate and taxable value coverage. The figure on the right shows the proportion of districts in our final sample that have property tax rate data in a given year. Data were hand collected from state Department of Revenue and Department of Education Annual Reports.

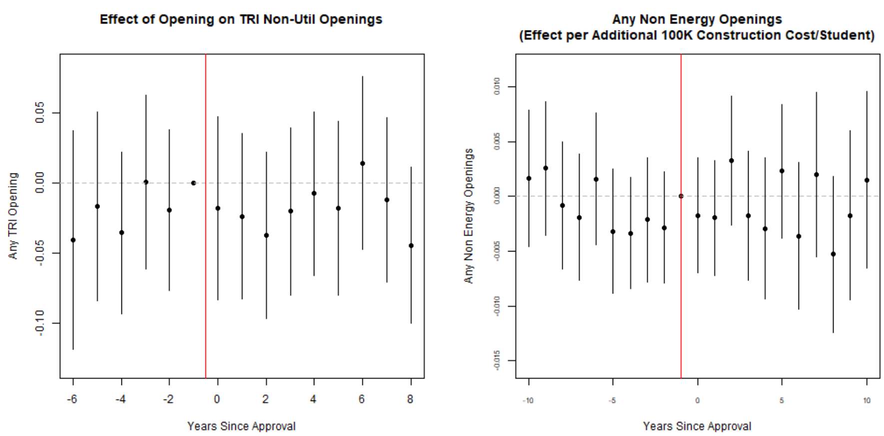


Figure A.3: Opening of Non-Utility TRI Facilities

This figure shows the likelihood of a district having any facility opening in a year surrounding the start of construction on a power plant. Opening data comes from the Toxic Release Inventory (TRI) and is based on the first year that a facility appears in the data. Coefficients come from a regression of the relevant outcome variable on district fixed-effects, year x border pair fixed effects and interactions between indicators for years since approval (-1 is the omitted category) and whether or not a district receives a plant. Estimated effect of the opening comes from dividing estimated construction costs (created using parameters provided in the EIA's Annual Energy Outlook) divided by the number of students in a district at the year of approval. Border pairs are any two districts that share a border within the same state and only one of the two districts experienced a plant opening between 1995 and 2015. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally. Plant data comes from the Toxic Release Inventory run by the EPA. Plant opening data comes from the Energy Information Administration (EIA).

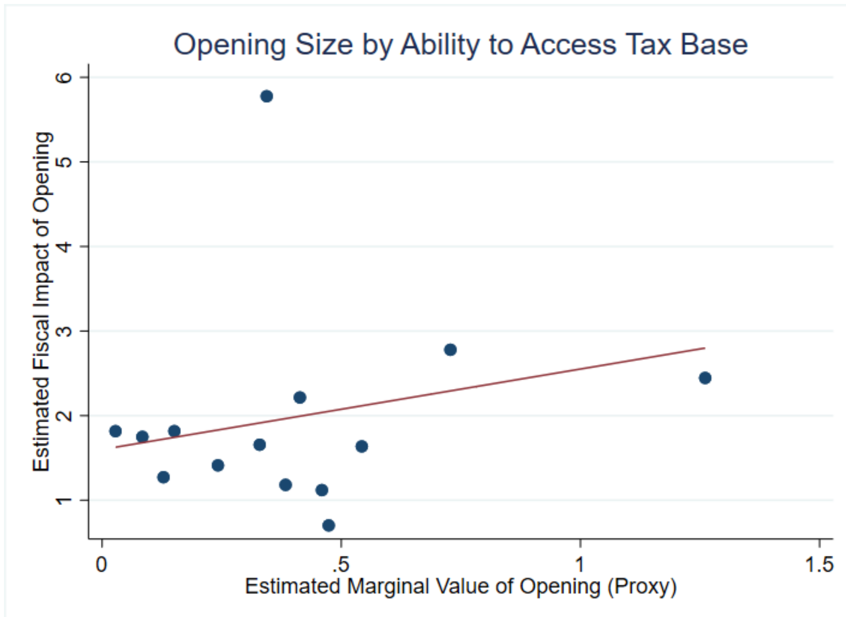


Figure A.4: Estimated Tax Base Effect of Opening by Estimated Marginal Value of Tax Base in Opening State-Year

This figure shows the relationship between the expected marginal value of an additional dollar of tax base with respect to school spending and the estimated tax base impact of a plant. Estimated tax base impact of a plant is equal to the plant's estimated construction cost divided by the number of students in a district in the year of approval. Construction costs were estimated using parameters from the EIA's Annual Energy Outlook. We proxy for district's estimated marginal value of tax base using a coefficient dervied from a state-specific regression of state and local revenue per student on taxable value per student with district and year fixed-effects.

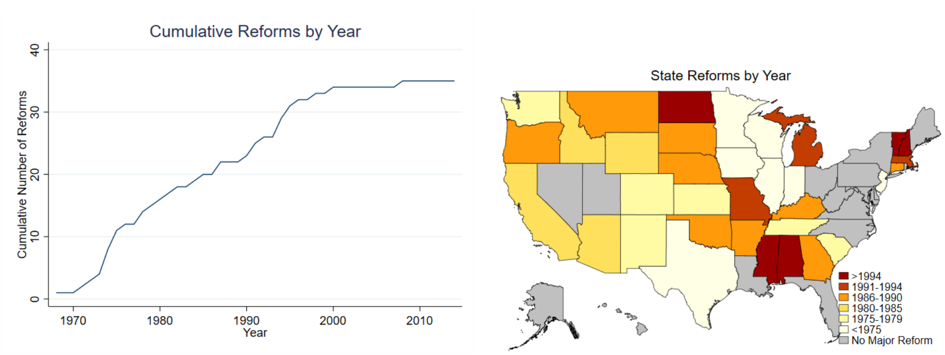
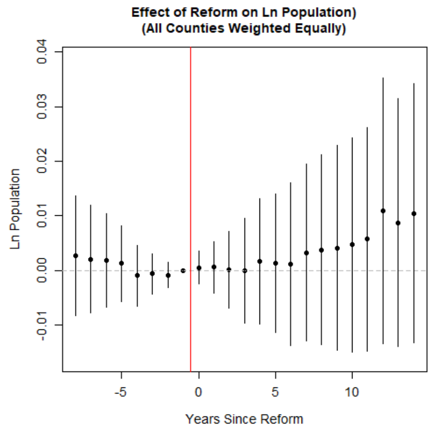


Figure A.5: School Finance Reforms Geographic and Temporal Distribution

This figure shows the cumulative number of school finance reforms affecting the marginal value of an additional dollar of tax base (left panel) and their geographic distribution (right panel). Reforms were identified using funding formulas reported in the Public School Finance Programs of the United States series.



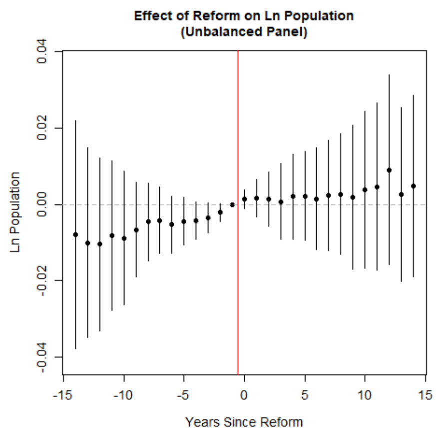


Figure A.6: Effect of School Finance Reform on Ln County Population

This figure shows the effect of a school finance reform on large local manufacturing establishments (>500 employees) and manufacturing employment per 1,000 population using a contiguous border county difference-in-differences design. Coefficients come from a regression of the outcome variable on indicators for years since the reform. Controls include border pair by year fixed effects and county fixed-effects. We cluster standard errors at the state and state border pair level. County-years with outcomes greater than the 99th percentile were excluded as outliers. Counties with less than 5,000 population in 1970 were also excluded. All employment and establishment data come from County Business Patterns (CBP).

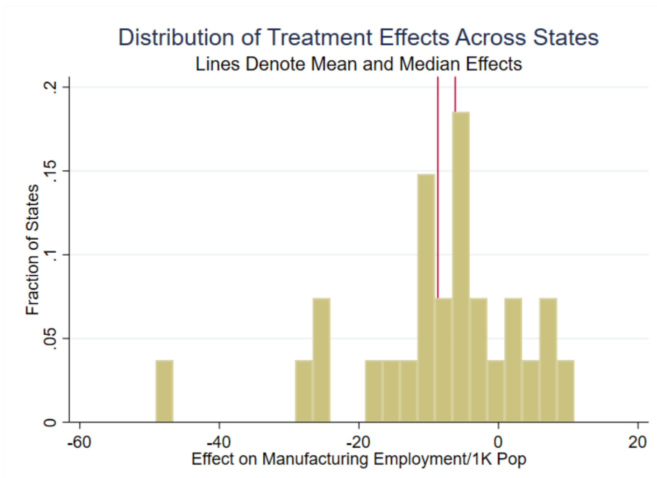


Figure A.7: Distribution of Treatment Effects by Reform State

This figure shows the effect of a school finance reform on manufacturing employment per 1,000 population using a contiguous border county difference-in-differences design. Coefficients come from a regression of the outcome variable on indicators for years since the reform run separately for each treatment state. Controls include border pair by year fixed effects and county fixed-effects. We cluster standard errors at the state and state border pair level. County-years with outcomes greater than the 99th percentile were excluded as outliers. Counties with less than 5,000 population in 1970 were also excluded. All employment and establishment data come from County Business Patterns (CBP).

Table A.1: Correlates of Plant Opening

VARIABLES	(1) Ln Stud	(2) URM	(3) Pct FRL	(4) Loc Rev/Stud ('000)	(5) Tot Rev/Stud ('000)	(6) Ln 1990 Home Val
Treated District	0.0389	0.0222**	0.00855	0.141	0.115	0.00241
	(0.0585)	(0.00943)	(0.00802)	(0.120)	(0.150)	(0.0168)
Observations	2,660	2,660	2,372	2,624	2,626	2,392
R^2	0.731	0.845	0.786	0.804	0.780	0.877

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the relationship between various economic, demographic and school funding variables and the probability that a district ever has a plant locate within it. All demographic characteristics come from the year a plant opens and are taken from the National Center of Education Statistics (NCES).

Table A.2: Effects of Plant Opening on School District Tax Base, Revenue and Expenditure Outcomes: Including Covariate by Year FE

VARIABLES	(1) Ln Tax Base /Stud	(2) Tax Rate (Mills)	(3) Loc Rev /Stud	(4) Tot Rev /Stud	(5) Tot Exp /Stud
Treat x Post Yrs 0-2	0.0210	0.136	0.196***	0.157*	0.144
	(0.0130)	(0.130)	(0.0706)	(0.0869)	(0.231)
Treat x Post Yrs 3-8	0.111***	-0.115	0.581***	0.450***	0.731*
	(0.0267)	(0.153)	(0.133)	(0.128)	(0.381)
Observations	19,508	19,290	34,824	34,824	35,704
R^2	0.976	0.977	0.967	0.955	0.885
Pair x Year FE	Y	Y	Y	Y	Y
Sample	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small
Dep. Var. Mean	580887	10.56	4.991	11.79	12.15

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the natural log of taxable value per student. Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state (unless otherwise noted) and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. . Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). All specifications also include baseline student, baseline free lunch share and 1990 home value by year by state fixed effects. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally unless otherwise noted. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table A.3: Effects of Plant Opening on School District Tax Base, Revenue and Expenditure Outcomes: Unbalanced Panel

VARIABLES	(1) Ln Tax Base /Stud	(2) Tax Rate (Mills)	(3) Loc Rev /Stud	(4) Tot Rev /Stud	(5) Tot Exp /Stud
Treat x Post Yrs 0-2	0.0189*	0.177**	0.150***	0.133**	0.109
	(0.0111)	(0.0737)	(0.0509)	(0.0585)	(0.137)
Treat x Post Yrs 3-8	0.104***	-0.0758	0.441***	0.345***	0.546**
	(0.0191)	(0.108)	(0.0918)	(0.0914)	(0.244)
Observations	47,798	47,292	83,150	83,000	84,752
R^2	0.970	0.973	0.949	0.927	0.846
Pair x Year FE	Y	Y	Y	Y	Y
Sample	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small
Dep. Var. Mean	580887	10.56	4.991	11.79	12.15

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the natural log of taxable value per student. Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state (unless otherwise noted) and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally unless otherwise noted. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table A.4: Effects of Plant Opening on School District Tax Base, Revenue and Expenditure Outcomes: All Openings Included

VARIABLES	(1) Ln Tax Base /Stud	(2) Tax Rate (Mills)	(3) Loc Rev /Stud	(4) Tot Rev /Stud	(5) Tot Exp /Stud
Treat x Post Yrs 0-2	0.0219*	0.0343	0.121**	0.0939	0.0222
	(0.0122)	(0.107)	(0.0543)	(0.0739)	(0.155)
Treat x Post Yrs 3-8	0.0911***	-0.244*	0.404***	0.329***	0.604**
	(0.0222)	(0.127)	(0.0998)	(0.104)	(0.303)
Observations	24,090	23,942	45,414	45,334	46,332
R^2	0.966	0.975	0.956	0.937	0.855
Pair x Year FE	Y	Y	Y	Y	Y
Sample	All	All	All	All	All
Dep. Var. Mean	584616	10.57	5.007	11.84	12.22

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the natural log of taxable value per student. Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state (unless otherwise noted) and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. . All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally unless otherwise noted. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table A.5: Effects of Plant Opening on School District Tax Base, Revenue and Expenditure Outcomes: By Plant Type

VARIABLES	(1) Tax Base /Stud	(2) Tax Base /Stud	(3) Rate (Mills)	(4) Rate (Mills)	(5) Loc Rev /Stud	(6) Loc Rev /Stud	(7) Tot Rev /Stud	(8) Tot Rev /Stud
Treat x Post Yrs 0-2	0.00818	0.0481**	-0.0107	0.245	0.0970	0.215*	0.0982	0.199
	(0.0168)	(0.0200)	(0.154)	(0.167)	(0.0651)	(0.119)	(0.0967)	(0.137)
Treat x Post Yrs 3-8	0.0701**	0.158***	-0.170	-0.166	0.358***	0.752***	0.226**	0.727***
	(0.0303)	(0.0356)	(0.190)	(0.189)	(0.116)	(0.226)	(0.113)	(0.240)
Observations	13,668	7,572	13,060	8,032	26,952	12,032	26,892	12,070
R^2	0.958	0.975	0.979	0.964	0.963	0.940	0.936	0.925
Pair x Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Sample	NG	WND	NG	WND	NG	WND	NG	WND
Dep. Var. Mean	496473	660553	10.69	10.48	4.608	5.469	11.33	12.86

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the natural log of taxable value per student. Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant and whether or not the district receives a plant. Border pairs are any two districts that share a border within the same state (unless otherwise noted) and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. . Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). Effects for natural gas plants are reported in the odd columns and wind turbines are reported in the even columns. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally unless otherwise noted. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table A.6: Effects of Plant Opening on School District Tax Base, Revenue and Expenditure Outcomes: Triple Difference Design

VARIABLES	(1) Ln Tax Base	(2) Tax Rate	(3) Loc Rev	(4) Tot Rev	(5) Tot Exp
Treat x Post Yrs 0-2 x Ln Exp Tax Base/Stud	0.0111	0.241***	0.0177**	0.00540	0.00454
	(0.00885)	(0.0693)	(0.00733)	(0.00390)	(0.00554)
Treat x Post Yrs 3-8 x Ln Exp Tax Base/Stud	0.0431***	0.239***	0.0571***	0.0135***	0.0212***
	(0.0163)	(0.0868)	(0.0110)	(0.00494)	(0.00741)
Observations	24,076	23,936	45,410	45,332	46,332
R^2	0.974	0.978	0.975	0.955	0.909
Pair x Year FE	Y	Y	Y	Y	Y
Sample	Non-Small	Non-Small	Non-Small	Non-Small	Non-Small
Dep. Var. Mean	584453	10.40	4.924	11.68	12.22

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the natural log of taxable value per student. Coefficients come from a regression of log taxable value/student on district fixed-effects, year x border pair fixed effects and interactions between periods relative to the treatment district of the pair gaining approval for a plant, whether or not the district receives a plant and the log expected tax base impact per student. Expected tax base per student is calculated by Border pairs are any two districts that share a border within the same state (unless otherwise noted) and only one of the two districts experienced a plant opening between 1995 and 2015. Only openings with data 6 years prior to an opening and 8 years following an opening are included unless otherwise indicated. . Only openings with greater than \$10,000 in expected tax base per student are included ($\approx$ 10% of openings are dropped). Estimated effect of the opening comes from dividing estimated construction costs (created using parameters provided in the EIA's Annual Energy Outlook) divided by the number of students in a district at the year of approval. All district-years are weighted by the inverse of the total number of districts attached to a given opening in a year in order to weight all plant openings equally unless otherwise noted. We implement two-way clustered standard errors at the plant (all border districts attached to a given opening) and district level. District property tax rates were hand-collected from state Department of Education and Department of Revenue's annual reports. Local revenues per student came from the National Center for Education Statistics (NCES). Plant opening data comes from the Energy Information Administration (EIA).

Table A.7: Effects of Plant Opening on Hedonic Characteristics of Homes Sold

VARIABLES	(1) Lot Size	(2) Lot Size	(3) Bedrooms	(4) Bedrooms	(5) Age	(6) Age	(7) Sqft	(8) Sqft	(9) SFH	(10) SFH
Treat x Post Yrs 0-2	-0.00992	-0.0253	0.0370*	0.0291	-0.457	0.249	6.598	12.50	0.00379	3.85e-05
	(0.0122)	(0.0211)	(0.0194)	(0.0437)	(0.382)	(0.713)	(5.801)	(11.22)	(0.00527)	(0.00709)
Treat x Post Yrs 3-8	-0.00992	-0.0215	0.0137	-0.0509*	-0.759	-0.741	-0.0295	-0.000517	0.000648	-0.00517
	(0.00849)	(0.0166)	(0.0205)	(0.0297)	(0.494)	(0.482)	(4.915)	(10.17)	(0.00524)	(0.00777)
Observations	387,584	133,615	402,197	100,502	441,313	140,445	492,910	189,787	501,701	198,578
R^2	0.890	0.899	0.542	0.580	0.845	0.865	0.932	0.894	0.822	0.828
Bandwidth	2 km	2 km	2 km	2 km	2 km	2 km	2 km	2 km	2 km	2 km
Weighted	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	All	No CA	All	No CA	All	No CA	All	No CA	All	No CA
Dep Var Mean.	0.383	0.383	3.203	3.203	20.86	20.86	1736	1736	0.829	0.829

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on various hedonic characteristics using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and .004 degree latitude x .004 degree longitude x year fixed-effects. Different observations for each outcome occur because different variables have different levels of missing values in the Zillow database. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—values of the outcome variables above the 99th percentile are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than $10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table A.8: Effects of Plant Openings on Home Prices: Repeat Sales Only

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price
Treat x Post Yrs 0-2	0.0246	0.0246	0.0141	0.0224	0.0348	0.0141
	(0.0172)	(0.0172)	(0.0168)	(0.0169)	(0.0341)	(0.0164)
Treat x Post Yrs 3-8	0.0293	0.0293	0.0161	0.0310*	0.0500	0.0180
	(0.0180)	(0.0180)	(0.0148)	(0.0169)	(0.0342)	(0.0150)
Observations	308,590	308,590	192,032	324,121	125,988	198,133
R^2	0.875	0.875	0.942	0.881	0.827	0.941
Weighted	Y	Y	Y	Y	Y	Y
Spec	Repeat-sales	Repeat-sales	Repeat-sales	Repeat-sales	Repeat-sales	Repeat-sales
Distr Sample	>10K Pop	>10k Pop	>10k Pop	All	All	All
State Sample	All	No-CA	CA-Only	All	No-CA	CA-Only

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and an individual parcel fixed-effect. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. . Openings with an expected tax base impact per student of less than $10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table A.9: Effects of Plant Opening on Quantity of Homes Sold

VARIABLES	(1) Ttl Sales	(2) Any Sales	(3) New Sales	(4) Any New	(5) Old Sales	(6) Any Old Sales
Treat x Post Yrs 0-2	13.39	0.0301	5.925	0.0381	7.463	0.0329
	(12.81)	(0.0257)	(7.488)	(0.0344)	(6.574)	(0.0251)
Treat x Post Yrs 3-8	22.49*	0.000793	8.503	0.0851**	13.99*	0.00610
	(12.71)	(0.0268)	(5.695)	(0.0362)	(8.383)	(0.0255)
Observations	3,268	3,268	3,268	3,268	3,268	3,268
R^2	0.944	0.722	0.796	0.862	0.966	0.724
Weighting	All Plant =	All Plant =	All Plant =	All Plant =	All Plant =	All Plant =
Dep. Var. Mean	81.38	0.831	15.36	0.342	66.01	0.818

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on the quantity of homes sold using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year fixed-effects and border pair by district fixed-effects. Standard errors are clustered at the plant district level. Sample restrictions are as indicated. All housing data come from the Zillow ZTRAX database—sales below $5,000 and above$1,500,000 as likely outliers. Only openings with an expected tax base impact of more than $10,000/student are included.

Table A.10: Effects of Plant Openings on Home Prices: Excluding New Construction

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price
Treat x Post Yrs 0-2	0.0310**	0.0543**	0.00605	0.0351**	0.0615**	0.00449
	(0.0137)	(0.0219)	(0.0115)	(0.0153)	(0.0252)	(0.0111)
Treat x Post Yrs 3-8	0.0399***	0.0700***	0.00979	0.0469***	0.0776***	0.0137
	(0.0132)	(0.0186)	(0.0149)	(0.0153)	(0.0238)	(0.0149)
Observations	409,654	164,804	244,850	436,812	184,493	252,319
R^2	0.751	0.692	0.827	0.766	0.728	0.830
Weighted	Y	Y	Y	Y	Y	Y
Distr Sample	>10K Pop	>10k Pop	>10k Pop	All	All	All
State Sample	All	No-CA	CA-Only	All	No-CA	CA-Only

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border and that were not constructed within a year of sale. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and a border pair by district by .004 degree latitude and .004 degree longitude fixed-effect. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than $10,000 are excluded as are any homes sold in their construction year. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table A.11: Effects of Plant Openings on Home Prices: Different Sample Criteria and Fixed-Effect Models

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price	(7) Ln PRice
Treat x Post Yrs 0-2	0.0217	0.0247	0.00775	0.0233	0.0208	0.0283	0.0467**
	(0.0327)	(0.0187)	(0.0203)	(0.0153)	(0.0150)	(0.0229)	(0.0192)
Treat x Post Yrs 3-8	0.0673**	0.0615***	0.0294	0.0468***	0.0304**	0.0373*	0.0550***
	(0.0286)	(0.0179)	(0.0231)	(0.0158)	(0.0139)	(0.0199)	(0.0188)
Constant	-0.0196**	0.00307	0.0179***	0.00293	0.00730**	-0.00577	-0.0150***
	(0.00776)	(0.00457)	(0.00618)	(0.00372)	(0.00362)	(0.00526)	(0.00454)
Observations	201,810	375,062	378,741	502,930	487,046	504,227	504,677
R^2	0.730	0.743	0.774	0.676	0.789	0.584	0.651
Weighted	Y	Y	Y	Y	Y	Y	Y
Spec	Brdr >5km from Plnt	No Cnty Bndry	Balanced	.008 Dgr FE	.001 Deg FE	Distr FE	Border Pair x Yr FE

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and .004 degree latitude x .004 degree longitude x year fixed-effects. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. "Brdr >5km" means all transactions in the border pair are at least 5km from the plant. "No Cnty Bndry" means there is no county boundaries within 2.5 km of the school district boundary. Balanced means that all openings have at least data for at least 6 years before approval and 8 years after approval. ".008 Dgr FE" means that a .008 degree latitude x .008 degree longitude x border pair x district fixed-effect is included as the geographic fixed-effect. '.001 Dgr FE" means that a .001 degree latitude x .001 degree longitude x border pair x district fixed-effect is included as the geographic fixed-effect. "Distr FE" means that a school district x border pair x district fixed-effect is included as the geographic fixed-effect. "Border pair by year" means border pair by year (instead of border pair by year by month) is included as the time varying fixed-effect, a .004 degr x .004 degr x border pair by district fixed-effect is also included as in the main specification. All housing data come from the Zillow ZTRAX database—sales below $5,000 or greater than $1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than $10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table A.12: Effects of Plant Openings on Home Prices: By Plant Type

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price	(7) Ln Price	(8) Ln Price
Treat x Post Yrs 0-2	0.0281*	-0.0219	0.0281*	-0.0219	0.0273**	0.0371	0.0273**	0.0371
	(0.0142)	(0.0615)	(0.0142)	(0.0615)	(0.0136)	(0.0691)	(0.0136)	(0.0691)
Treat x Post Yrs 3-8	0.0398**	0.123***	0.0398**	0.123***	0.0406**	0.122*	0.0406**	0.122*
	(0.0151)	(0.0377)	(0.0151)	(0.0377)	(0.0172)	(0.0667)	(0.0172)	(0.0667)
Observations	478,916	22,783	478,916	22,783	512,032	26,321	512,032	26,321
R^2	0.701	0.730	0.701	0.730	0.712	0.757	0.712	0.757
Weighted	Y	Y	Y	Y	Y	Y	Y	Y
Spec	.4km grid	.4km grid	.4km grid	.4km grid	.4km grid	.4km grid	.4km grid	.4km grid
Dist Sample	.>10K Pop	.>10K Pop	>10K Pop	.>10K Pop	All	All	All	All
Sample	NG	REN	NG-No CA	REN-No CA	NG	REN	NG-No CA	REN-No CA

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on log local housing prices using a border difference-in-difference design. We only include sales within a bandwidth of 2 km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year by month fixed-effects and a border pair by district by .004 degree latitude and .004 degree longitude fixed-effect. The outcome variable is residualized for hedonic by state fixed-effects, which include land-use , home age by plant district (5 year bins with 1 year bins for ages <5), bedrooms , bathrooms, square footage (500 sq ft bins), heating type and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Each observation is weighted by the inverse of the number of sales in its treated unit each year (i.e. border pair x treat). Standard errors are clustered at the plant district level. All housing data come from the Zillow ZTRAX database—sales below \$5,000 or greater than \$1,500,000 are excluded as outliers. Only pairs of districts in which both district have a 2000 population greater than 10,000 are included in the regression. Openings with an expected tax base impact per student of less than \$10,000 are excluded. Data are from fourteen states in total: Arizona, California, Colorado, Georgia, Illinois, Indiana, Iowa, Michigan, Minnesota, New York, North Carolina, Ohio, Oklahoma and Pennsylvania. These make up roughly 60% of openings. Texas and Kansas both have large number of openings, but do not have publicly available home sale data.

Table A.13: Effects of Plant Opening on Key School Finance Variables: Home Price Analysis

VARIABLES	(1) Loc Rev	(2) Loc Rev	(3) Tot Rev	(4) Tot Rev	(5) Ttl Exp	(6) Ttl Exp
Treat x Post Yrs 0-2	0.281***	0.338***	0.331	0.464***	0.273	0.0116
	(0.105)	(0.109)	(0.234)	(0.167)	(0.361)	(0.418)
Treat x Post Yrs 3-8	0.311	0.484***	0.247	0.579***	0.284	0.684
	(0.200)	(0.139)	(0.241)	(0.203)	(0.344)	(0.471)
Observations	498,086	197,591	497,589	197,365	498,086	197,591
R^2	0.978	0.991	0.958	0.989	0.923	0.934
Cut-off	>10K	>20K	>10K	>20K	>10K	>20K
Weighted	Y	Y	Y	Y	Y	Y
Sample	All	No CA	All	No CA	All	No CA

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on key school finance variables using a border difference-in-differences design. We only include sales within a bandwidth of 2km from the border. Coefficients come from a regression of log home prices on an interaction between an indicator for whether a house is in a plant-receiving district and a vector of indicators for grouping of years since approval. Controls include border pair by year fixed effects and border pair by district x distance to border (200 m bins) x distance to plant (200 m bins) fixed effects. Standard errors are clustered at the plant district level. Sample restrictions are as indicated. Only openings with an expected tax base impact of more than $10,000/student are included.

Table A.14: Effects of Plant Opening on Nearby Home Prices: Robustness Check

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price	(5) Ln Price	(6) Ln Price
5-10KM x Post Yrs 0-2	0.00210		0.00344		-0.00702	
	(0.0140)		(0.0174)		(0.0167)	
5-10KM x Post Yrs 3-8	-0.0289		-0.0102		-0.0354	
	(0.0210)		(0.0194)		(0.0240)	
$<$5km x Post Yrs 0-2	0.00192		-0.0106		-0.0309*	
	(0.0143)		(0.0150)		(0.0176)	
$<$5km x Post Yrs 3-8	-0.0630**		-0.0388		-0.0679**	
	(0.0299)		(0.0246)		(0.0302)	
Ln Dist x Post Yrs 0-2		-0.00673		0.0375*		0.0148
		(0.0140)		(0.0205)		(0.0130)
Ln Dist x Post Yrs 3-8		0.0422*		0.0478**		0.0291
		(0.0226)		(0.0191)		(0.0194)
Observations	1,764,776	1,764,776	1,487,430	1,487,430	1,441,980	1,441,980
R^2	0.703	0.703	0.721	0.721	0.850	0.850
Size Cutoff	$>$100MW	$>$100MW	$>$100MW	$>$100MW	$>$100MW	$>$100MW
Hedonics	Y	Y	Y	Y	Y	Y
Max Dist	15km	15km	20km	20km	20km	20km
Weighted	Y	Y	Y	Y	Y	Y
Model	Base	Base	Balanced Panel	Balanced Panel	Repeat Sale	Repeat Sale

Clustered standard errors in parentheses
*** p$<$0.01, ** p$<$0.05, * p$<$0.1

This table shows the effect of a plant opening on local housing prices for nearby homes. All regressions include plant district by year, .004 degree longitude x .004 degree latitude bins, and district x year fixed-effects unless otherwise indicated.The outcome variable is residualized for hedonic by state controls. These include land-use , home age by plant district (5 year bins with 1 year bins for ages $<$5), bedrooms, square footage (500 sq ft bins), and lot size (1 acre bins). Missing hedonics are included as a separate indicator. Only homes within 20km of the opening plant are included unless otherwise indicated. Only plants who are in districts with a population greater than 10,000 are included in weighted analysis. Standard errors are clustered at the plant district level. Only openings with at least two years of pre and two years of post data are included except for the balanced panel specification in which only plants with at least 6 years of pre data and 8 years of post data are included. All housing data come from the Zillow ZTRAX database—sales that are less than \$5,000 or greater than \$1,500,000 are excluded as outliers. Weighted specifications are weighted by the inverse of the number of sales within 20km of the plants in a given year.

Table A.15: Effects of Plant Opening on Nearby Home Prices: No District by Year FE

VARIABLES	(1) Ln Price	(2) Ln Price	(3) Ln Price	(4) Ln Price
5-10Km x Post Yrs 0-2	0.00423			
	(0.0107)			
5-10Km x Post Yrs 3-8	-0.00804			
	(0.0128)			
<5Km x Post Yrs 0-2	0.0163			
	(0.0126)			
<5Km x Post Yrs 3-8	-0.00596			
	(0.0213)			
Ln Dist x Post Yrs 0-2		-0.00872		
		(0.00946)		
Ln Dist x Post Yrs 3-8		0.000391		
		(0.0126)		
5-10Km x Post Yrs 0-2 x Nameplate Capac ('00s MW)			0.000140	
			(0.00215)	
5-10Km x Post Yrs 3-8 x Nameplate Capac ('00s MW)			-0.00551**	
			(0.00265)	
<5Km x Post Yrs 0-2 x Nameplate Capac ('00s MW)			-0.000511	
			(0.00282)	
<5Km x Post Yrs 3-8 x Nameplate Capac ('00s MW)			-0.00760**	
			(0.00374)	
Ln Dist x Post Yrs 0-2 x Nameplate Capac ('00s MW)				0.000949
				(0.00197)
Ln Dist x Post Yrs 3-8 x Nameplate Capac ('00s MW)				0.00673**
				(0.00277)
Observations	2,228,460	2,228,460	3,964,622	4,107,596
R^2	0.677	0.677	0.684	0.670
Size Cutoff	>100MW	>100MW	N	N
Hedonics	Y	Y	Y	Y
Max Dist	20km	20km	20km	20km
Weighted	Y	Y	Y	Y

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a plant opening on local housing prices for nearby homes. All regressions include plant district by year and .004 degree longitude x .004 degree latitude bins.The outcome variable is residualized for hedonic by state controls. Only homes within 20km of the opening plant with at least two years of pre and two years of post data are are included. Standard errors are clustered at the plant district level.

Table A.16: Census Tract Demographics by Distance to Plant and District Status

VARIABLES Type	(1) LnInc Absolute	(2) LnInc Pctile	(3) LnHomeVal Absolute	(4) LnHomeVal Pctile	(5) OwnShare Absolute	(6) OwnShare Pctile	(7) WhiteShare Absolute	(8) WhiteShare Pctile
In Distr, $\leq$5KM from Plant	-0.0557***	-0.0510***	-0.0552***	-0.0426**	-0.0534***	-0.134***	-0.0311***	-0.0938***
	(0.0164)	(0.0181)	(0.0191)	(0.0202)	(0.00853)	(0.0176)	(0.00940)	(0.0199)
In Distr, 5KM-15KM from Plant	-0.00516	-0.0118	0.00270	0.00792	-0.0390***	-0.0919***	-0.0237***	-0.0776***
	(0.0127)	(0.0143)	(0.0151)	(0.0162)	(0.00668)	(0.0146)	(0.00778)	(0.0166)
Outside Distr,$\leq$5KM from Plant	-0.0660***	-0.0542**	-0.0879***	-0.0578**	-0.0473***	-0.0787***	-0.0394***	-0.0707***
	(0.0225)	(0.0221)	(0.0248)	(0.0243)	(0.0114)	(0.0213)	(0.0129)	(0.0231)
Outside Distr, $>$15KM from Plant	-0.0203	-0.0345**	-0.0610***	-0.0332**	-0.0459***	-0.0869***	-0.0345***	-0.0622***
	(0.0147)	(0.0149)	(0.0189)	(0.0169)	(0.00694)	(0.0147)	(0.00894)	(0.0170)
Observations	1,894	1,879	1,893	1,879	1,894	1,879	1,894	1,879
R^2	0.714	0.203	0.831	0.202	0.470	0.222	0.832	0.221

Twoway clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the association between being within a given distance of a plant opening and inside or outside the plant's district and various demographic outcomes. Demographic outcomes categorized are absolute show the association with that demographic's value and outcomes categorized as percentiles shows the association between a tract's outcome's percentile within their school district. The omitted variable are tracts that are inside the school disrict but more than 15km from the plant. All regressions control for plant district fixed-effects and weight all plant districts equally. Demographic data come from the 2000 Census.

Table A.17: School Finance Reform and Power Plants: Openings and Retirements

VARIABLES	(1) Open 50MW	(2) Open 50MW	(3) Open 250MW	(4) Open 250MW	(5) Retire 50MW	(6) Retire 50MW	(7) Retire 250MW	(8) Retire 250MW
Treat x Post Yr $\leq$5	-0.00348	-0.00510*	-0.00323	-0.00368	0.00217	0.000969	0.00174	0.001000
	(0.00226)	(0.00288)	(0.00212)	(0.00283)	(0.00157)	(0.00170)	(0.00141)	(0.00158)
Treat x Post Yr >5	-0.00174	-0.00159	-0.000473	-2.83e-07	0.000196	-0.000630	0.000454	-0.000504
	(0.00129)	(0.00148)	(0.000987)	(0.00120)	(0.00150)	(0.00154)	(0.00145)	(0.00154)
Observations	212,448	122,938	212,448	122,938	212,448	122,938	212,448	122,938
R^2	0.524	0.522	0.521	0.520	0.545	0.556	0.527	0.533
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	0.00395	0.00368	0.00245	0.00245	0.00165	0.00147	0.00102	0.00102

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on power plant openings and retirements. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event.

Table A.18: Baseline Differences in Key Demographic and Economic Characteristics

VARIABLES	(1) Urban Share	(2) Urban Share	(3) Ln Pop	(4) Ln Pop	(5) White Share	(6) White Share	(7) Pov Share	(8) Pv Share
Treat	0.00918	0.000126	-0.0731	-0.140	0.00107	0.00200	-0.00799	-0.00413
	(0.0165)	(0.0208)	(0.0764)	(0.105)	(0.00584)	(0.00662)	(0.00482)	(0.00640)
Observations	9,000	5,350	9,000	5,350	9,000	5,350	9,000	5,350
R^2	0.608	0.594	0.730	0.709	0.856	0.862	0.805	0.810
Balanced Panel	N	Y	Y	N	Y	N	Y	N

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on baseline covariates Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on an indicator for whether a district is in a reform state. Controls include county pair fixed effects. Standard errors are clustered at the state and state border pair level. All data are from the US Census. Population data are annual estimates, while other outcomes are assigned the value of the most recent Decennial Census. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event.

Table A.19: Baseline Differences in Manufacturing and Power Plant Exposure

VARIABLES	(1) EmpManf	(2) EmpManf	(3) BigManfEst	(4) BigManfEst	(5) AnyPlant	(6) AnyPlant	(7) AnyLrgPlnt	(8) AnyLrgPlnt
Treat	11.73***	12.24**	0.0126***	0.0137**	0.0125	0.00747	0.0150	0.00663
	(3.596)	(5.031)	(0.00417)	(0.00544)	(0.0247)	(0.0368)	(0.0186)	(0.0283)
Observations	8,112	4,256	8,768	4,748	9,034	4,894	9,034	5,382
R^2	0.686	0.689	0.654	0.657	0.549	0.551	0.526	0.511
Balanced Panel	Y	N	Y	N	Y	N	Y	N
Sample	Base	Base	Base	Base				

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on baseline covariates Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on an indicator for whether a district is in a reform state. Controls include county pair fixed effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event.

Table A.20: School Finance Reform and Manufacturing Employment: Robustness Check

VARIABLES	(1) LnEmpManf	(2) LnEmpManf	(3) EmpManf	(4) EmpManf	(5) EmpManf	(6) EmpManf	(7) EmpManf	(8) EmpManf	(9) EmpManf	(10) EmpManf
Treat x Post Yr ≤5	-0.0575**	-0.0606**	-3.287**	-2.569	-3.020**	-2.290	-1.533	-1.731	-3.316*	-2.880
	(0.0270)	(0.0287)	(1.571)	(1.672)	(1.447)	(1.398)	(2.347)	(1.826)	(1.817)	(2.182)
Treat x Post Yr >5	-0.113***	-0.0946**	-7.270***	-5.910**	-8.112***	-7.026***	-4.965**	-4.878**	-7.753**	-6.355*
	(0.0372)	(0.0409)	(2.517)	(2.727)	(1.930)	(2.116)	(2.308)	(2.145)	(3.064)	(3.324)
Observations	165,576	91,792	156,840	83,012	56,618	29,860	37,976	32,756	89,902	63,312
R^2	0.953	0.957	0.956	0.961	0.957	0.967	0.954	0.961	0.963	0.965
Robustness	Log DV	Log DV	Pre-1998	Pre-1998	w/i .5 ln(bl manuf)	w/i .5 ln(bl manuf)	w/i 25/1K	w/i 25/1K	No outlier states	No outlier states
Dep. Var Mean	70.29	70.29	72.01	72.01	85.46	85.46	56.23	56.23	61.96	61.96

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Sample restrictions are as indicated. "Outlier states" refers to states whose baseline difference in manufacturing employment was greater than 20 workers per 1,000 population. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.21: School Finance Reform and Manufacturing Establishments: Robustness Check

VARIABLES	(1) LnLrgManfEst	(2) LnLrgManfEst	(3) LrgManfEst	(4) LrgManfEst	(5) LrgManfEst	(6) LrgManfEst	(7) LrgManfEst	(8) LrgManfEst	(9) LrgManfEst	(10) LrgManfEst
Treat x Post Yr ≤ 5	-0.117**	-0.109**	-0.00680**	-0.00495*	-0.00492**	-0.00423	-0.00324	-0.00407	-0.00306*	-0.00359**
	(0.0497)	(0.0480)	(0.00259)	(0.00250)	(0.00238)	(0.00260)	(0.00378)	(0.00281)	(0.00159)	(0.00176)
Treat x Post Yr >5	-0.169***	-0.147***	-0.0118***	-0.00959***	-0.0108***	-0.0104***	-0.00682*	-0.00929***	-0.00718**	-0.00630**
	(0.0533)	(0.0542)	(0.00354)	(0.00348)	(0.00272)	(0.00313)	(0.00367)	(0.00301)	(0.00284)	(0.00302)
Observations	201,752	117,162	190,660	106,154	59,498	31,992	45,546	94,870	116,132	85,332
R^2	0.902	0.912	0.900	0.915	0.900	0.916	0.889	0.910	0.910	0.916
rob1ustness	Log DV	Log DV	Pre-1998	Pre-1998	w/i .5 ln(bl manuf	w/i .25 ln(bl manuf	w/i .25 ln(bl manuf	w/i .25 ln(bl manuf	States w/ no dif	States w/ no dif
Dep. Var Mean	68.48	68.48	0.0608	0.0608	0.0743	0.0743	0.0461	0.0630	0.0484	0.0484

Twoway clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Sample restrictions are as indicated. "Outlier states" refers to states whose baseline difference in manufacturing employment was greater than 20 workers per 1,000 population. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.22: School Finance Reform and Employment and Establishments for Non-Manufacturing Industries

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Non-Manf	Non-Manf	Non-Manf Est>250Emp	Non-Manf Est>250Emp	Non-Manf Est>500Emp	Non-Manf Est>500Emp
Treat x Post Yrs $\leq$5	0.0329	0.0611	-6.72e-05	0.000609	-0.000407	-0.00141
	(1.317)	(1.562)	(0.00100)	(0.00114)	(0.000652)	(0.00128)
Treat x Post Yrs ¿5	0.115	0.364	0.000876	0.00243	0.000211	-0.000402
	(2.066)	(2.498)	(0.00177)	(0.00190)	(0.000750)	(0.00122)
Observations	167,412	92,432	208,062	120,052	119,938	119,938
R^2	0.970	0.973	0.873	0.885	0.879	0.892
Weight	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	186.3	160.7	0.0277	0.0324	0.0104	0.0134

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on employment outcomes by industry using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Employment data come from County Business Patterns and are suppressed for counties with few establishments. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. We exclude outcome values greater than the 99th percentile as outliers.

Table A.23: School Finance Reform and Employment by Industry Type

VARIABLES	(1) AgEmp/1KPop	(2) MineEmp/1KPop	(3) ConstrEmp/1KPop	(4) RtlEmp/1KPopr	(5) TransUtilEmp/1KPop	(6) WholesaleEmp/1KPop	(7) OthInd/1KPop
Treat x Post Yr $\leq$5	0.0859	-1.667	0.108	-0.0799	-0.472	0.251	0.207
	(0.171)	(1.223)	(0.240)	(0.531)	(0.286)	(0.259)	(1.344)
Treat x Post Yr >5	0.184	-0.751	-0.238	-0.464	-0.562	0.439	0.0101
	(0.205)	(1.521)	(0.344)	(0.791)	(0.432)	(0.398)	(2.124)
Observations	24,686	20,780	102,058	118,298	87,842	103,678	120,498
R^2	0.859	0.931	0.913	0.954	0.911	0.935	0.922
Dep. Var Mean	1.612	11.27	11.65	47.08	11.17	13.60	85.82

)
Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on employment outcomes by industry using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Employment data come from County Business Patterns and are suppressed for counties with few establishments. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. We exclude outcome values greater than the 99th percentile as outliers. There are different numbers of observations across differnet industries because many smaller industries have many county-year employment information suppressed to preserve privacy.

Table A.24: School Finance Reform and Power Plant Openings: School District Overlap

VARIABLES	(1) Open>25MW	(2) Open>25MW	(3) Open>25MW	(4) Open>25MW	(5) Open>100MW	(6) Open>100MW	(7) Open>100MW	(8) Open>100MW
Treat x Post Yr $\leq$5 x Overlap	-7.399*	-6.682	0.00492	0.00641	-0.0210	-0.0274	-0.0233	-0.0373
	(4.294)	(5.723)	(0.00847)	(0.00736)	(0.0272)	(0.0445)	(0.0300)	(0.0402)
Treat x Post Yr $\leq$5 x Overlap	-24.61***	-29.17***	-0.00733	-0.0115	-0.0400	-0.0498	-0.0556	-0.0716
	(8.596)	(10.45)	(0.00948)	(0.0117)	(0.0471)	(0.0650)	(0.0429)	(0.0535)
Observations	197,854	114,286	206,366	119,352	212,112	122,708	212,112	122,708
R^2	0.950	0.954	0.912	0.922	0.972	0.976	0.968	0.971
Dep. Var Mean	69.30	69.30	0.0609	0.0609	0.216	0.216	0.140	0.140

)
Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design by the degree of overlap between a counties' school and zoning jurisdictions. Overlap is measured as the interaction between population-weighted HHI of school districts within each zoning jurisdiction in a county and the population-weighted HHI of zoning districts within each school district in a county. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.25: School Finance Reform and Power Plant Openings: SUTVA Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	EmpManf	EmpManf	LrgMnfEst	LrgMnfEst	AnyGen(50)	AnyGen(50)	AnyGen(250)	AnyGen(250)
Treat x Post Yrs $\leq$5 x Dist to Bord.	0.0325		3.17e-05		0.000107		5.63e-05	
	(0.0720)		(0.000130)		(0.000611)		(0.000519)	
Treat x Post Yrs >5 x Dist to Bord.	-0.0304		-3.65e-05		0.000194		0.000229	
	(0.0870)		(0.000142)		(0.000705)		(0.000662)	
Treat x Post Yrs $\leq$5		-1.596		-0.00353		-0.0101	-0.0152	-0.0105
		(1.675)		(0.00231)		(0.0155)		(0.0151)
Treat x Post Yrs >5		-7.008**		-0.00778**		-0.0294	-0.0322	-0.0242
		(2.919)		(0.00345)		(0.0190)		(0.0196)
Observations	176,266	95,792	232,134	121,578	238,364	125,856	238,364	125,856
R^2	0.957	0.958	0.915	0.902	0.970	0.964	0.964	0.953
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced	Balanced
Sample Type	All	>50 Mi border	All	>50 Mi border	All	>50 Mi border	All	>50 Mi border
Dep. Var Mean	70.51	67.24	0.0598	0.0529	0.227	0.224	0.133	0.133

Twoway clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 90 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers. Distance to border refers to the mininum distance of the control in a pair to its nearest treated neighbor.

Table A.26: School Finance Reform and Power Plant Openings: Weighting by State

VARIABLES	(1) EmpManf	(2) EmpManf	(3) LrgMnfEst	(4) LrgMnfEst	(5) AnyPlnt50	(6) AnyPlnt50	(7) AnyPlnt500	(8) AnyPlnt500
Treat x Post Yr $\leq$5	-4.699***	-3.589**	-0.00618***	-0.00638**	0.0165	-0.00108	0.00632	-0.00957
	(1.659)	(1.726)	(0.00217)	(0.00237)	(0.0154)	(0.0158)	(0.0113)	(0.0104)
Treat x Post Yr >5	-10.33***	-9.210***	-0.0121***	-0.0128***	-0.0233	-0.0297	-0.0284	-0.0331
	(2.468)	(2.799)	(0.00325)	(0.00397)	(0.0236)	(0.0254)	(0.0208)	(0.0224)
Observations	165,876	94,038	210,828	121,876	212,448	122,938	212,448	122,938
R^2	0.959	0.965	0.911	0.925	0.960	0.964	0.956	0.960
Weight	All Treated State=	All Treated State=	All Treated State=	All Treated State=	All Treated State=	All Treated State=	All Treated State=	All Treated State=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	68.20	67.79	0.0539	0.0532	0.234	0.215	0.159	0.144

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Observations are weighted such that all treated states count equally. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.27: School Finance Reform and Power Plant Openings: Weighting by Population

VARIABLES	(1) EmpManf	(2) EmpManf	(3) LrgMnfEst	(4) LrgMnfEst	(5) AnyPlnt50	(6) AnyPlnt50	(7) AnyPlnt500	(8) AnyPlnt500
Treat x Post Yr $\leq$5	-5.651***	-3.240*	-0.00734**	-0.00720**	-0.0678	-0.116	-0.0686	-0.115
	(1.559)	(1.851)	(0.00291)	(0.00318)	(0.0528)	(0.0862)	(0.0528)	(0.0862)
Treat x Post Yr >5	-10.41***	-9.654**	-0.00981***	-0.00973***	-0.222**	-0.298**	-0.222**	-0.293**
	(3.333)	(4.033)	(0.00349)	(0.00323)	(0.0957)	(0.120)	(0.0952)	(0.122)
Observations	165,876	91,940	210,828	121,876	212,448	122,938	212,448	122,938
R^2	0.964	0.968	0.927	0.938	0.947	0.933	0.943	0.928
Weight	Pop.	Pop.	Pop.	Pop.	Pop.	Pop.	Pop.	Pop.
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Dep. Var Mean	93.70	92.01	0.0760	0.0744	0.497	0.493	0.390	0.393

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Observations are weighted by treated county population times the inverse of the number of county pairs a given treated county has. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.28: School Finance Reform and Manufacturing Employment by Bandwidth

VARIABLES	(1) EmpManf	(2) EmpManf	(3) EmpManf	(4) EmpManf	(5) EmpManf	(6) EmpManf	(7) EmpManf	(8) EmpManf	(9) EmpManf	(10) EmpManf
Treat x Post Yr ≤5	-5.068**	-4.357**	-3.021**	-2.008	-2.752*	-2.051	-1.876	-1.264	-1.529	-0.846
	(2.218)	(1.979)	(1.486)	(1.378)	(1.455)	(1.527)	(1.350)	(1.425)	(1.301)	(1.378)
Treat x Post Yr >5	-7.561***	-8.129***	-6.879***	-6.244***	-7.027***	-6.131***	-5.500***	-4.586**	-5.111**	-4.267**
	(2.714)	(2.710)	(2.073)	(2.233)	(2.056)	(2.226)	(1.877)	(2.093)	(1.921)	(2.110)
Observations	20,950	11,338	69,724	38,962	161,216	89,556	291,016	159,870	446,176	247,654
R^2	0.956	0.962	0.957	0.962	0.954	0.959	0.952	0.956	0.952	0.956
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Max BW	15	15	45	45	60	60	75	75	90	90
Dep. Var Mean	75.25	78.45	71.82	74.95	69.36	73.18	68.98	72.72	68.20	71.98

Twoway standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. County pairs whose geographic centroids were greater than the max bandwidth specified were excluded in each analysis. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.29: School Finance Reform and Manufacturing Establishments by Bandwidth

VARIABLES	(1) LrgMnfEst	(2) LrgMnfEst	(3) LrgMnfEst	(4) LrgMnfEst	(5) LrgMnfEst	(6) LrgMnfEst	(7) LrgMnfEst	(8) LrgMnfEst	(9) LrgMnfEst	(10) LrgMnfEst
Treat x Post Yr $\leq$5	-0.0116***	-0.00648*	-0.00692**	-0.00329	-0.00659**	-0.00482	-0.00599**	-0.00514*	-0.00574**	-0.00469
	(0.00416)	(0.00326)	(0.00310)	(0.00265)	(0.00275)	(0.00290)	(0.00288)	(0.00293)	(0.00272)	(0.00287)
Treat x Post Yr >5	-0.0143***	-0.00942*	-0.0116***	-0.00884**	-0.0115***	-0.00936***	-0.00974***	-0.00864***	-0.00860***	-0.00737**
	(0.00512)	(0.00466)	(0.00358)	(0.00347)	(0.00313)	(0.00327)	(0.00290)	(0.00301)	(0.00284)	(0.00297)
Observations	27,216	15,380	69,724	43,134	161,216	98,716	291,016	176,758	446,176	272,004
R^2	0.902	0.917	0.902	0.922	0.903	0.918	0.900	0.913	0.900	0.913
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Max BW	15	15	45	45	60	60	75	75	90	90
Dep. Var Mean	0.0698	0.0711	0.0665	0.0713	0.0649	0.0682	0.0636	0.0678	0.0627	0.0667

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing outcomes using a county border pair difference-in-differences design. County pairs whose geographic centroids were greater than the max bandwidth specified were excluded in each analysis. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers.

Table A.30: School Finance Reform and Power Plant Openings by Bandwidth

VARIABLES	(1) AnyGen	(2) AnyGen	(3) AnyGen	(4) AnyGen	(5) AnyGen	(6) AnyGen	(7) AnyGen	(8) AnyGen	(9) AnyGen	(10) AnyGen
Treat x Post Yr $\leq$5	-0.0102	-0.00259	-0.00281	-0.00292	-0.00583	-0.0110	-0.00160	-0.00881	-0.00106	-0.00956
	(0.0134)	(0.0135)	(0.00909)	(0.00930)	(0.00892)	(0.00979)	(0.00813)	(0.00796)	(0.00712)	(0.00868)
Treat x Post Yr >5	-0.0152	-0.0130	-0.0150	-0.0141	-0.0281*	-0.0302*	-0.0200*	-0.0237*	-0.0173	-0.0230*
	(0.0151)	(0.0182)	(0.0135)	(0.0159)	(0.0141)	(0.0164)	(0.0119)	(0.0135)	(0.0106)	(0.0128)
Observations	27,438	15,838	91,978	54,032	212,448	122,938	384,308	220,086	588,130	338,396
R^2	0.970	0.973	0.968	0.972	0.965	0.969	0.966	0.971	0.965	0.970
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Max BW	30	30	45	45	60	60	75	75	90	90
Dep. Var Mean	0.246	0.233	0.222	0.217	0.214	0.215	0.211	0.212	0.217	0.220

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were closer than the indicated maximum bandwidth were included in each regression. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event.

Table A.31: School Finance Reform and Manufacturing Employment: Alternate Reform Identification Strategy

VARIABLES	(1) EmpManf	(2) EmpManf	(3) LrgManf	(4) LrgManf	(5) AnyGen50	(6) AnyGen50	(7) AnyGen500	(8) AnyGen500
Treat x Post Yrs ≤5	-2.596*	-2.112	-0.00654**	-0.00586*	-0.0245**	-0.0466**	-0.0137	-0.0300
	(1.300)	(1.626)	(0.00244)	(0.00293)	(0.00954)	(0.0186)	(0.00848)	(0.0197)
Treat x Post Yrs >5	-4.972*	-3.841	-0.00928***	-0.00657*	-0.0365**	-0.0685**	-0.0140	-0.0393
	(2.570)	(3.310)	(0.00327)	(0.00365)	(0.0171)	(0.0259)	(0.0156)	(0.0260)
Observations	179,022	74,092	215,728	79,856	229,150	86,710	229,150	86,710
R^2	0.955	0.959	0.899	0.913	0.966	0.965	0.959	0.958
Weight	All County=	All County=	All County=	All County=	All County=	All County=	All County=	All County=
Controls	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year	Pair x Year
Sample	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced	Unbalanced	Balanced
Dep. Var Mean	74.12	84.93	0.0671	0.0671	0.224	0.224	0.148	0.148

Twoway clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a school finance reform on manufacturing and power generation outcomes using a county border pair difference-in-differences design. Counties whose geographic centroids were less than 60 miles apart were included. Coefficients come from a regression of the outcome variable on the interaction between an indicator for whether a district is a reform state and a vector of indicators for periods relative to reform (pre-reform is the omitted category). Controls include county pair by year fixed effects and county fixed-effects. Standard errors are clustered at the state and state border pair level. Manufacturing data come from County Business Patterns. Power plant data come from EIA Form 860. Only plants that were owned by utilities still in operation in 1990 were included. Counties with fewer than 1,000 population in 1970 are excluded. Our unbalanced sample consists of 14 years before and after the reform year, while our balanced sample consists of 8 years prior and 14 years following the event. Outcomes with values greater than the 99th percentile are excluded as outliers. Reform events were the first event in each state identified by Jackson, Johnson, and Persico (2014).

A.2 ZTRAX Database

Home sales data come from the Zillow ZTRAX database and are merged to assessment records using parcel ID. We restrict sales to properties categorized by Zillow as single-family units. This includes single-family homes and condominiums, but excludes multi-family units although results are robust to their inclusion. Included sales are non-foreclosures with a deed type that does not reflect a transfer between family members, an inheritance, or another non-market transfer of property. These sample restrictions are designed to capture arm's length transactions. Foreclosures are transactions flagged by Zillow as foreclosures, as well as tax deeds, foreclosure deeds, commissioner's deeds, redemption deeds, deeds in lieu of foreclosure, receiver's deeds, sheriff's deeds, beneficiary deeds, notices of sale, and notices of lease pendens. This is a liberal definition of foreclosure that includes the first notice of foreclosure.

For all sale types, we assume that a house will only transact once in a 93 day window.[41] We define a transaction event as beginning with the first time a parcel transacts. If another transaction is recorded within the next 93 days, that transaction is considered part of the initial transaction, and we check for another transaction within the following 93 days, until a 93 day period with no transaction activity passes.[42] The transaction date is coded as the date of the first event. The price is the maximum price observed over the transaction window.

Transaction and assessment data come from county governments. Because data are provided at the county level, the years in which counties enter our sample differ even within a state. However, we cannot simply use the first year a county has a transaction in the data as the year in which data becomes available for a county because many counties include a small minority of transactions ($< .1\%$ of housing units) for many years in the past before reporting all transactions. Accordingly, we identify the starting year for each county in the following way.

[41]Many transaction records only provide a month and year of sale. The 93 day window allows for any three month window regardless of month length.

[42]Many events have multiple transactions recorded in the ZTRAX database due to mortgage changes, adjustments, multiple foreclosure notices, etc.

We first identify all years in which a county had a greater than 300% increase in sales (off a minimum of a base of 5 transactions). This threshold is chosen because it is greater than any increase we would expect to observe in the course of normal annual fluctuations and therefore is likely driven by changes in reporting. We then define a county's initial year as the most recent year in which there was a greater than 300% increase observed in our data (or the first year transactions are recorded if $>$300% increase never occurred.) We drop all transactions prior to our empirically-defined "start-year" from our analysis. Results are robust to alternative specifications of start year. Finally, because our home price analysis uses a border difference-in-differences design that in some cases span counties we drop all transactions within the border-pair prior to the year in which the last county began full transaction reporting to ensure sample consistency.

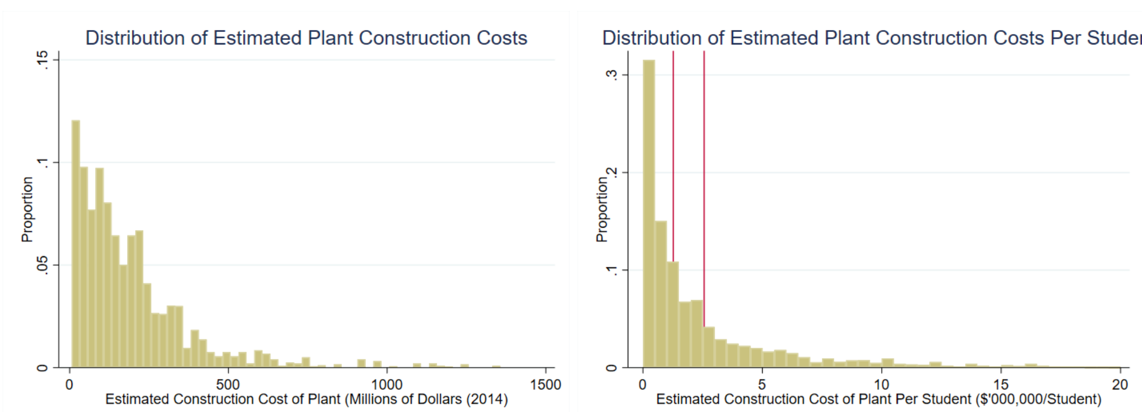
A.3 Imputing Plant Value

To proxy for the effect of a plant opening on the local tax base, we use the estimated overnight construction cost of a power plant. Overnight construction cost is a term of art that reflects the estimated hypothetical cost of building a power plant overnight so as to abstract away from borrowing costs. To do this, we used annual estimates of overnight construction costs per kWh taken from the EIA's Annual Energy Outlook between 1997 and 2018. For years 1995 and 1996 we used the 1997 values. For combined-cycle gas turbines and combustion gas turbines, values for basic and advanced turbines were given. We averaged these two values for each year, but results are robust to using either one. All estimates were adjusted for inflation and are presented in 2014 dollars.

An important note is that EIA estimates are presented for the construction of a power plant of a given size. We use the resulting cost per kwh for all power plants. If, as is likely, economies of scale exist then we are understating the costs of small plants and overstating the costs of large plants. This would bias our results toward zero and so to the extent that this affects our overall results they should be thought of as a lower-bound.

The left panel of Figure A.8 shows the distribution of estimated plant construction costs, while the right panel shows the distribution of costs per student. In both cases the distribution is right-skewed; there is a very long right tail of expected impacts. To allow for better interpret ability, we drop all impacts above 2 million dollars of construction costs per student (<2% of plants) from the figure. Most districts have increases in expected tax base large enough to expect meaningful fiscal impact for local schools—the median opening has roughly $127,000 in estimated construction costs per student, while the mean opening has approximately $256,000 per student in estimated construction costs.[43]

Figure A.8: Estimated Fiscal Impact



This figure shows various summary statistics on estimated plant construction costs, which we use as a proxy of plant valuation. These figures show the distribution of estimated plant construction costs and construction costs per student (we exclude the small number of plant openings with grater than $2 million in construction costs per student). Construction cost data and plant opening data come from the EIA, while enrollment data comes from the NCES.

[43]Most school district property taxes are between .3%-1.5%.

A.4 Identifying Reforms

In Section 4, we estimate how shocking a district's marginal value of tax base with respect to school spending affected local land-use decisions. To do this, we identified school finance legislation, litigation and initiatives that affected this quantity within a given state between 1970 and 2015. To identify these reforms we used information from *Public School Finance Programs in the United States* 1962, 1967, 1972, 1976, 1979, 1994, 1998, 2007, 2011, 2015, and 2018. Broadly speaking, changes to the marginal value of tax base are determined by the extent to which increases in local revenue crowd-out state and federal transfers and the level of taxes a local district can charge. Accordingly, in each report year we attempted to quantify a state's school funding formula and tax limitations. We then looked for major changes in crowd-out or tax limits between report years and identified these changes as potential reforms. We next turned to the text of the report and online searches to identify the legislation, litigation or initiatives that led to these changes in order to ensure that such a change had indeed occurred and to identify the year in which the reform took place. If a state had multiple reform years, we used the first reform year only. Below, we summarize the year, reform type and changes of each reform used in our analysis.

Reforms are typically not amenable to simple summary statistics (except for example, where no crowd-out exists and so total crowd-out for all districts is 0). In the table below, we provide information on the level of crowd-out as a function of a district's property value (P). This is for an "average" district in a state, but should not be thought of as holding for all districts. For example, some districts with high property values may generate more money than their foundation value through local sources and so face no effective crowd-out when increasing the tax base. When necessary, we attempt to provide additional context. We also attempt to describe the tax limitations in place for each state. These should not be compared across states as assessment ratios (the ratio of assessed value to true market value) changed dramatically over time. Unfortunately, for earlier years we often lack information on the assessment ratio used and so cannot inflate these rates into a value common across all states.

Figure A.9 shows the correlation between our qualitatively-identified reforms and the first major reform identified by Jackson, Johnson, and Persico (2014) in their paper examining the long-run educational and labor market effects of these reforms. Note that Jackson, Johnson, and Persico (2014) were searching for reforms that changed the distribution of school funding, not events that shocked the marginal value of tax base per se, and so therefore we would not necessarily expect to identify all the same reform events. Each state in our reform sample also appeared in their sample and there is a fairly strong correlation; roughly 3/5 of the reforms identified in our sample occurred within three years of reforms identified by Jackson, Johnson, and Persico (2014). There were a further ten states that had reforms as identified by Jackson, Johnson, and Persico (2014), but for which we did not find a major change in their marginal value of tax base or tax limitations. Nonetheless, this coincidence in reforms is reassuring.

Figure A.9: Correlation Between Qualitatively-Defined Reforms and Reforms Identified in Jackson, Johnson, and Persico (2014)

This figure compares the year of our qualitatively-identified reform with the reform in a state identified by Jackson, Johnson, and Persico (2014).

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
AL	1995	Response to litigation	.0025*(1938 Ttl State Property Value/Ttl State Property Value)*P	.01*P	15 mills (exceed w/ vote)	15 mills (exceed w/ vote)	Crowd-out
AZ	1980	Legislation	0	.0472*P	None	State-set expenditure limit	Crowd-out
AR	1984	Litigation	Equalization aid roughly 10% of state aid and distributed as function of AV/Teacher rank	.025*P	None	State-set expenditure limit	Crowd-out
CA	1978	Response to litigation + Prop 13 (property tax limitation)	.0387*P	≈full crowd-out	No limit (w/ vote)	.01 (total taxes)	Crowd-out + Limitation
CO	1974	Legislation	min(.017,tax needed to raise $250/student). Tax needed to raise most common, effective crowd-out of 0	Guaranteed Revenue Base of $29,620 (almost full crowd-out up to this limit—binds for most districts)	No limit (w/ vote)	No limit (w/vote)	Crowd-out

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post re-form tax limit	Shock Driver
CT	1990	Legislation	Guaranteed Tax Base in which each district was guaranteed $\tau^* P_8 5$, but b/c underfunded only received pro-rated share of available funding (which was very low)	τ^* $Inc_D/IncMax$ where $\tau*$ is the tax necessary to create the foundation amount at the state guaranteed wealth level (1.567 median wealth)	No limit	No limit	Crowd-out
DE	None						
FL	None						
GA	1986	Legislation	Tax rate necessary to raise 78,000,000 on all property in state (only 10% of state funding, so rate likely low)	.00825*P (if wealth <90th percentile), otherwise .005*P		20 (no limit w/ vote)	Crowd-out
ID	1979	Initiative	.022*P	.0036*P	.027 (no limit with vote)	..004	Limitation
IL	1973	Legislation	1.12*.011*P	Guaranteed tax base of $42,000 (or more if elem or hs distr) for any tax rate (affects most >80% of districts)	No limit w/ vote	No limit w/ vote	Crowd-out

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
IN	1973	Legislation	.0215*P	.03*P	.047	min(.03, tax to generate last year levy)	Crowd-out+Limitation
IA	1972	Legislation	0	0054*P	No limit w/vote	109% last year's levy	Crowd-out+Limitation
KS	1975	Legislation	.005*County Valuation (apportioned by district share of county employees)	.017*P	No limit w/vote	107% last year's levy	Crowd-out
KY	1991	Litigation	0	.0036*P	Unclear	Unclear	Crowd-out
LA	None						
ME	None						
MD	None						
MA	1993	Legislation	Complicated, but >75% of districts (1979) in hold-harmless and therefore had effectively 0 crowd-out	.0094*P*Inc_d /AvgInc, but complicated so does not apply uniformly to all districts	.025 (all rates)	.025 (all rates)	Crowd-out
MI	1993	Initiative	τ*P if value/student <$40K (up to first 30 mills)	.018*P	.05 (all taxes)	.021	Crowd-out (Proportional) + Limitation

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
MN	1971	Litigation	.019*P (up to foundation leve, but foundation low—1/3 off formula)	.03*P (foundation level dramatically increased)	None	No limit w/ vote	Crowd-out + Limitation
MS	1994	Legislation	0	.028*P (up to 28% of program cost)	10% increase	.055	Crowd-out
MO	1993	Legislation	Complicated function in income, that leads to little crowd-out	τ*P* Inc/AvgInc	No-limit w/ vote	No limit w/vote	Crowd-out
MT	1989	Litigation	0	(τ^*)*P where $\tau*$ is the tax rate necessary to get 40% of state funding on 175% x avg state val (so 0 If >175% Avg Wealth)	No w/ vote	No w/ vote	Crowd-out
NE	1990	Litigation	Each district's share is .012*P and then receive pro-rated share from eq aid available (but typically aid quite low—i.e. 1979 $70/pupil on avg, so true crowd-out negligible)	.0124*P (but bites, eq aid now 5x as large as 1979)	No w/ vote	Levy to collect 3%-5% > previous year	Crowd-Out + Limitation
NV	None		.				

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
NH	1998	Litigation	Funds 8% of edu cost adjusted by wealth (in reality funding much lower, so almost all districts face effective crowd-out of 0)	avg state tax rate*P for districts with <Avg Wealth (.0237 in 2007), 0 for >Avg Wealth	None	None	Crowd-out
NJ	1970	Legislation	.0105*P	Guaranteed Tax Base of $30K, so full crowd-out if less and none if more	No-limit	No limit	Crowd-out
NM	1974	Legislation	.0005*P	.0089*P	.002		Crowd-Out + Limitation
NY	None						
NC	None						
ND	2007	Legislation	.035*P	Guaranteed tax base equal to 180 mills at 90% average wealth/stud and recapture above 150% avg wealth/stud	.185	.185	Crowd-out
OH	None						
OK	1990	.018*P (but many districts (≈50% held harmless)	Nearly full crowd-out for districts with <$55K/student				

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
OR	1991	Initiative	τ^{*}*P where τ^{*} is tax rate chosen to distribute equalization dollars (typically quite low)	Nearly full crowd-out	No limit w/ vote	.005 (1991)	Crowd-out+Limitation
PA	None						
RI	None						
SC	1977	Legislation	None	τ^{*}*P where τ^{*} is tax rate necessary to generate 30% of guaranteed funding	No	No	Crowd-out
SD	Legislation	1986	.018*P	τ^{*}*P where τ^{*} is avg. non-agri tax rate (came into effect in 1990)	.053	.0186	Crowd-out + Limitation
TN	1977	Legislation	0 (use index of econ. ability that does not include prop wealth)	τ^{*}*P where τ^{*} is amount necessary to raise 10% of guaranteed funding in avg district	No limit	No limit	Crowd-out
TX	1973	Legislation	0 (use index of econ. ability that does not include prop wealth)	.003*P	.015	.015	Crowd-out

St	Yr	Type	Pre-Reform Crowd-out	Post-Reform Crowd-out	Pre-reform tax limit	Post reform tax limit	Shock Driver
UT	None						
VT	1997	Legis	.01297*P (but many districts off formula)	Full crowd-out (began in 1997 and effectively full state funding in 2001)	no	NA	Crowd-out + Limitation
VA	None						
WA	1974	Legislation	.0119*P	All property tax rev sent to state w/ exception of small (25% of budget) local option levies	No limit w/ vote	.015	Crowd-out + Limitation
WV	None						
WI	1973	Legis	τ^{*}*P where τ^{*} is tax producing equalization amt in district with 45k val/stud	τ^{*}*P where τ^{*} is tax producing ttl cost in district with 98k val/stud	.025	.025	Crowd-out
WY	1983	Litigation	.01*P	.025*P (recapture if above 109% foundation guarantee)	.028	.028	Crowd-out

Chapter 2

Enforcing Compliance: The Case of Automatic License Suspensions

2.1 Abstract

Non-incarcerative punishments such as fines are an essential part of the United States criminal justice system. Theoretically, the deterrence and distributional effects of these punishments will depend upon the consequences for punishment noncompliance. In this paper, I test this idea empirically using a unique natural experiment in Washington that first eliminated and then reinstated driver's license suspensions for traffic offense punishment noncompliance, a common state policy affecting millions of drivers annually. Mandating suspensions caused large increases in compliance, fine-repayment, and total punishment with greater effects for lower-income individuals, while also leading to suggestive declines in traffic accidents among low-income drivers.[1]

2.2 Introduction

In the United States, the vast majority of criminal offenses and traffic infractions involve non-incarcerative punishments. Unlike incarceration, compliance with these types of punishments is not compelled, but is instead an individual choice; after receiving their punishment, an offender can choose between complying with the terms of the punishment or facing the consequences of noncompliance. Thus, the actual expected punishment for many crimes is not the court-ordered punishment itself, but will instead be the lesser of the cost of complying with the punishment and the cost of whatever consequences exist for noncompliance. For instance, the expected cost of a speeding ticket in utility terms is not simply the cost of paying the ticket, but is instead the

[1] I thank the Washington Administrative Office of the Courts for dealing with my many questions. This work does not reflect the opinions of the Washington Administrative Office of the Courts.

minimum of the cost of paying the ticket and the cost of whatever will happen if the ticket is not paid (i.e. added financial penalties, automatic license suspensions, imprisonment, etc).

The existence of this compliance choice has two important implications. First, it suggests that the government can shift average punishment levels not only by changing statutory punishments, but also by changing the consequences of noncompliance. Second, in cases such as fines where punishment compliance costs are negatively correlated with income, a two-tiered punishment system can emerge; higher-income individuals will choose to accept the statutory punishment, while lower-income individuals will choose to face the consequences of noncompliance. As a result, different relative levels of statutory punishments and sanctions for punishment noncompliance that achieve the same level of average deterrence may have very different distributional consequences.

In this paper, I provide novel theoretical and empirical evidence on the average and distributional effects of changing consequences for noncompliance in a criminal justice context. Gaining a better understanding of these effects is particularly important because punishment noncompliance for low-level crimes and infractions is extremely common in the United States. Although no nationwide statistics exist, I find that 16% of traffic offenses in Washington state are out of compliance 18 months after the offense and that approximately 14% of all adults had their licenses suspended due to noncompliance at least once between 2008 and 2017. Legal scholars have found similar levels of noncompliance in other states; for instance, a report by the Legal Aid Center found that 10%-20% of licensed drivers in Virginia have their license suspended due to failure to pay a fine or appear in court (Ciaramella, 2016).

Thus, for a meaningful proportion of the population, and particularly the low-income population, the operative punishment for a criminal offense or traffic infraction is not the court-ordered punishment, but is instead whatever consequence exists for noncompliance. While a large body of work examines how changing statutory punishments affects crime and criminal offenders (Chalfin and McCrary, 2017), the effects of punishments for noncompliance have gone largely unstudied by economists. In this paper, I attempt to fill this gap by examining both theoretically and empirically how changing consequences for noncompliance with court-ordered punishments affects punishment compliance, overall punishment levels and long-run criminal behavior both on average and across the income distribution.

I begin by providing a simple theoretical framework that integrates an offender's punishment compliance decision into a model of deterrence. The model produces several key predictions: increasing costs of non-compliance should increase total punishment faced by low-income offenders, increase compliance by middle-income offenders and decrease crime rates by low-income individuals.

I next examine the model's predictions empirically using a natural experiment that caused a change in the consequence for noncompliance with traffic infractions and misdemeanors in Washington state. In June 2004, the Washington State Supreme Court invalidated a law mandating license suspensions for noncompliance with traffic offenses. In 2005, the Washington legislature responded by enacting a new law that reinstated these suspensions for any traffic violation that occurred on or after July 1st, 2005, creating a discontinuity in the consequences for noncompliance depending upon the timing of a violation. This law change was largely unpublicized[2] and I find no discontinuities around the law enactment date in the number of cases filed, case punishment severity or previous criminal history of offenders.

Accordingly, I begin by using a regression discontinuity design to test the effect of shifting to the license suspension regime on compliance and punishment outcomes. I show that this increase in the costs of noncompliance had a large effect on compliance (5 pp increase or 22%) and amount of fines repaid ($67 increase or 10%). This effect was largest for offenders from middle-income zip codes, consistent with the model's predictions. I next show that the new license suspension regime also greatly increased the total punishment faced by traffic offenders. Exposure to the policy increases the probability of a ever having license suspension for noncompliance

[2] I can find no mention of it in a newspaper archive search of major Washington newspapers at the time and there is no noticeable change in state-level Google trend search data for driver's licenses.

attached to a case by 24 percentage points for traffic misdemeanors and 6 percentage points for traffic infractions (relative to levels close to 0% for offenses committed prior to the policy change). These effects on punishment were much larger among individuals from the lowest-income zip codes. Suspensions also had additional legal ramifications including fines and jail time arising from Driving with License Suspended (DWLS) charges, which again were disproportionately large for residents of lower-income areas.

I next examine if the increase in the consequences for fine nonpayment caused changes in unsafe driving behavior as proxied through traffic accidents using a series of difference-in-difference and triple difference designs.[3] I exploit both the changes in consequences for noncompliance over time induced by the suspension policy being first eliminated and then reinstated, as well as the fact that changing the costs of complying with a fine should differentially affect individuals at the bottom of the income distribution. To control for secular trends in accident rates, I use Northern California as a control group.[4]

I begin by using vehicle age as a proxy for driver income and examine whether older vehicles were more likely to get in injury-producing accidents[5] during the period in which the Supreme Court decision prohibiting license suspensions was in effect in Washington relative to Northern California. I show that relative to Northern California, there was a marked increase in the average age of vehicles in injury-producing crashes in Washington state during the year in which the Supreme Court decision was active suggesting that low-income drivers were more likely to get into crashes when the costs of noncompliance fell. I then show that similar results hold when comparing changes in crashes per capita among high and low poverty jurisdictions in Washington state and Northern California—crash rates increased by 12% in high poverty

[3]Note that the offense-date cut-off used in the regression discontinuity design only creates variation in the noncompliance punishment for a given offense, but does not create variation in the noncompliance penalties for any future offenses, which is the object of interest in this analysis. Thus it is not possible to use a regression discontinuity design to examine the effects of the policy on overall deterrence.

[4]Northern California was chosen as a control group because it has similar geographic and economic characteristics to Washington state and has crash data available during the time period of the study.

[5]I also test robustness to only serious and fatal injuries to address endogenous reporting concerns.

jurisdictions relative to low-poverty jurisdictions when the Supreme Court decision was in effect.

Examining these analyses dynamically provides a much stronger than normal test of the parallel trends assumption because I can compare changes during this period to both the period prior to the court decision *and* to the period following the law change reinstating suspensions and I find no evidence for any violations. Results are also robust to the inclusion of a large number of time-varying controls, as well as other checks addressing endogenous reporting concerns. In sum, these results suggest that the license suspension policy is effective in increasing overall traffic safety, but does so in a highly regressive manner.

This paper contributes to both a theoretical and empirical literature on the economics of deterrence. In the years since the publication of the seminal Becker (1968) model of rational criminal-decisionmaking, an enormous literature has emerged extending his results both theoretically and empirically (Nagin, 2013; Chalfin and McCrary, 2017). To date, this work has largely focused on the direct deterrence effects of different types of statutory punishments, be it prison (Mueller-Smith, 2015; Aizer and Doyle Jr, 2015; Mueller-Smith and Schnepel, 2016; Bhuller et al., 2016), a blend of prison, rehabilitative classes, fines and license suspension (Hansen, 2015), speeding fines (Goncalves and Mello, 2017) or short-term license suspensions (Gehrsitz, 2017). In this paper, I extend this work by showing that in many cases it is not sufficient to examine the deterrence effect of the punishment alone, but instead is necessary to consider the interaction between statutory punishments and consequences for punishment noncompliance. This idea builds upon the theoretical work of Levitt (1997) and Polinsky (2006), both of whom create models that show when wealth is unobservable and fines are not costlessly enforced, a social planner will want to introduce imprisonment as part of a deterrence scheme. This paper extends their analysis by focusing on the average and distributional effects of changing each policy independently and importantly, provides empirical estimates of the effects of these changes.

This paper also contributes to a growing empirical literature on the costs and benefits of fines and forfeitures as an enforcement tool. Fines have been shown to deter traffic offenses

(Goncalves and Mello, 2017), but also impose large costs on affected individuals. Short-term (30 day) license suspensions have also been found to have negative effects on financial outcomes and lead to meaningful increases in borrowing and decreases in employment (Mello, 2018). Additionally, Makowsky, Stratmann, and Tabarrok (2018) show that when police agencies are the end recipients of money collected through fines and forfeitures, their policing decisions will be distorted towards revenue generating crimes. In this paper, I build on this literature by showing new theoretical and empirical evidence that noncompliance costs, an important consequence of fine-based systems of enforcement, differentially affect low-income offenders, but also deliver meaningful reductions in unsafe driving behavior. I also provide the first causal analysis of the most commonly-used noncompliance sanction, automatic license suspension, which affects millions of Americans each year.

The remainder of this paper is organized as follows. Section 2 provides descriptive evidence of noncompliance in the US criminal justice system. Section 3 provides a simple theoretical model that will guide my empirical analysis. Section 4 describes the policy setting and data used in the empirical analysis. Section 5 describes the empirical strategy for estimating the effect of greater compliance costs on individual compliance and punishment outcomes and then presents the results. Section 6 outlines the empirical strategy for the analysis of the effects of greater compliance costs on unsafe driving behavior and then presents the results. Section 7 concludes.

2.3 Background

Crimes that have non-incarcerative punishments make up the bulk of Americans' interactions with the criminal justice system. In 2016, there were 13.2 million misdemeanor cases filed, compared to less than 3 million felony cases (Stevenson and Mayson, 2018) and the number of traffic infractions per year is likely closer to 40 million.[6] Fines are a major component of the total punishment for most misdemeanors and traffic infractions.[7] Nonpayment of fines for traffic and criminal offenses is widespread across the United States, but consequences vary across states. The most common consequence is license suspension; approximately fifteen states suspend licenses for any unpaid court debt and more than 40 states impose license suspensions for unpaid debt related to traffic offenses. Although no systematic data exists on the total number of individuals with suspended licenses, state-level estimates have found that between 10%-20% of licensed drivers have a suspended license for failure to appear in court or pay a fine at any given time in a subset of states with strong suspension policies (Legal Aid Justice Center, 2018).

In this section, I will provide descriptive statistics on noncompliance with traffic offenses in Washington state, the setting of my empirical study. The descriptive facts in this section serve two purposes. First, they demonstrate that the types of policies studied in this paper have large impacts on a broad swath of the state's population—descriptive evidence that is largely lacking from the academic literature. Second, the trends discussed here will motivate how I theoretically model the effects of policies affecting punishment noncompliance.

I begin by showing the proportion of adults who were charged with a traffic infraction or misdemeanor between 2014 and 2016 in Washington state as a function of a zip code's median income (Figure 2.1, left panel).[8] The vast majority of these offenses are traffic infractions and speeding violations are the plurality offense. We first see that the proportion of adult with any traffic offense over this three-year period is quite high, above 25%. However, traffic offense rates are only weakly correlated with zip code income.[9]

[6]Although no nationwide estimates exist; the state of Washington, the setting of my study, has approximately 1 million traffic cases a year alone and makes up only 2% of the US population.

[7]Most felonies also carry large fines, but in a relative sense they are a smaller part of the overall punishment.

[8]I exclude all zip codes that have fewer than 1,000 people, are outside of Washington state or which do not match a 2000 Zip Code Tabulation Area (ZCTA). I use 2014-2016 because these are the years in which I observe the universe of all traffic infractions (I observe the universe of misdemeanors for beginning in 2003) with sufficient time to measure noncompliance.

[9]I provide results as shares of the population. This is not entirely correct as even within this three year period there will be some migration implying that the total potential population is larger than the population used here. However, given the three year timeframe, this difference is unlikely to be large.

I next show the proportion of individuals who have still failed to comply with the terms of their traffic offense as proxied by the issuance of a Failure to Appear (FTA) 18 months after their offense,[10] again by zip code median income (Figure 2.1, right-panel). Across all cases, 16% had at least one FTA notice active 18 months following the offense and a further 12% of cases had an already-resolved FTA over this time frame. These numbers suggest that the consequences of noncompliance are preferable to the costs of paying the fine for many offenders. Unsurprisingly, the income gradient in noncompliance is quite steep. Among zip codes with median income less than $40,000/year, nearly 25% of cases had one instance of noncompliance active at 18 months. Conversely, for zip codes with median income above $80,000, this number is closer to 5%. This relationship is not simply an artifact of different case composition among different income groups; even when I restrict individuals to have the same offense, the pattern remains.

I also provide descriptive evidence on the consequences of noncompliance. The left panel of Figure 2.2 shows the proportion of over-16 residents whose license was suspended for noncompliance with a punishment for an offense at any point over the ten years between 2008 and 2017. There is again a clear gradient by zip code income; in zip codes with less than $40,000 median income, roughly 20% of adult residents had a license suspension initiated at some point during this ten year time period. Conversely, for zip codes with greater than $80,000 median income, only about 7% of residents had a suspension. This number for low-income adults is remarkably high, highlighting the importance of understanding the costs and benefits of this policy.[11]

A suspended license has a potentially large effect on individual well-being. More than 70% of Washington workers drive alone to work and many individuals rely on vehicles to bring

[10]In the Washington data, FTAs signal failure to appear at a hearing, but also (and more commonly) for failure to comply with the terms of a punishment such as fine repayment.

[11]Note that this is an under-count of the actual number of individuals with a suspended license during this time period as many individuals will have had a suspension that began prior to 2008, but remained active. However, because we are now using a ten year period the number of potential individuals who could be suspended in each zip code will likely be significantly larger than the population because of migration so this may lead to an overestimate. Given the available data, I believe these are the best estimates possible.

children to school and run errands. Thus, the utility (and financial) costs of not driving are likely extremely high. However, the costs of driving with a suspended license are also high. If an individual is caught driving on a suspended license they can be arrested for driving with license suspended (DWLS), a class A misdemeanor. In general, the first offense is typically punished with large fines and a suspended jail sentence, but later offenses can lead to days or weeks in jail. The right panel of Figure 2.2 shows the proportion of adults with a DWLS charge between 2008 and 2017 by zip code median income. Individuals in zip codes with greater than $80,000 in median income have very low DWLS charge rates—around 1-2%—while in zip codes with less than $40,000 in median income nearly 10% of adults received such a charge during this ten year period.[12]

Together, these results show that noncompliance with traffic fines is a fairly common phenomenon, but occurs disproportionately among the poor. This noncompliance has large consequences; individuals lose access to their license and both lose the ability to legally drive to work or perform other necessary tasks in addition to becoming much more likely to be convicted of driving with a suspended license, an offense that carries large fines and the potential for jail time. In the next section, I use these facts as a guide while modeling how we expect changes to fines or consequences to fine nonpayment affect the overall and distribution of costs and benefits of criminal behavior.

2.4 Theoretical Framework

In this section, I introduce a simple conceptual framework for thinking about how changes to both fine levels and the consequences for nonpayment of fines affect total deterrence and the

[12]This relationship is even steeper than the gradient between incomes and suspensions suggesting that not only are lower income individuals more likely to have their licenses suspended, but conditional on suspension they are also more likely to have a DWLS charge. This may be because higher income individuals are less likely to be charged with crimes conditional on being caught, because higher-income drivers can more easily substitute away from driving, or because they have the means to hire higher quality lawyers conditional on being caught among other reasons.

distribution of costs across income groups. I assume individuals in a population have a level of disposable income that is uniformly distributed between 0 and 1. As in the standard Becker (1968) framework, the expected cost of committing a crime is:

$$C_j = f(p_j, f_j, u_j) \tag{2.1}$$

Where p_j is individual j's probability of being caught, f_j is his expected punishment if caught and u_j are other individual specific factors such as income that increase or decrease the utility loss caused by this punishment. For simplicity, I assume that the only individual-specific factor that affects punishment costs is an individual's income, i_j and that the probability of being caught, p_j, is independent of both income and expected punishment. I assume that $i_j \in (0,1)$ for all individuals. An individual can choose between two punishments, the statutory fine f_y or the consequence for fine nonpayment, f_n.[13] I assume that the utility costs of f_y are decreasing in income because of the diminishing marginal utility of income. Specifically, I let $C(f_y, i_j) = f_y(1 - i_j)$ where $f_y \in (0,1)$. For analytic simplicity I assume that the utility costs of f_n are constant across the population, where $C(f_n, i_j) = f_n$ and $f_n \in (0,1)$.[14] I further assume that $f_y > f_n$ so we are in the interesting case where some individuals will choose the fine and others will choose not to comply.[15]

I now examine how changing either f_y or f_n affects an individual's expected costs of punishment and the probability of committing a crime. To determine expected costs, we first

[13]Note that f_n groups together a nested set of compliance choices. For instance, if the consequence for nonpayment is a license suspension, an individual would then have to decide about whether or not he would comply with that suspension. I consider f_n to be a summary of the total costs that arise out of each of these compliance decisions. I assume this cost is also known to the potential offender using backwards induction.

[14]Note that for results to hold it will only be necessary that the relationship between non-monetary punishment and income be less negative than the relationship between the fine and income. Non-monetary punishments can be measured primarily in time—this is true for incarceration, but also for attending drug, alcohol and/or driving classes, license suspensions and even dealing with debt collection agencies. Higher income individuals have higher hourly wages and so time costs imply a larger income loss for these individuals (although of course income loss is only one dimension of the costs caused by non-monetary punishments). Because fines impose a constant monetary costs across individuals of all incomes, this suggests at the very least a less negative slope for the relationship between the costs of non-monetary punishments and income than the costs of fines and income.

[15]Under this set-up, if $f_n > f_y$, everyone would choose to comply with the fine.

describe an individual's punishment choice. An individual will choose the punishment that minimizes individual costs. Because the assumptions enumerated above imply that $C(f_y, i_j)$ and $C(f_n)$ cross at a single point, we know that for a given f_y and f_n there must exist some income level i^* such that:

- If $i_j > i^*$ then $C_j = C(f_y, i_j)$
- If $i_j < i^*$ then $C_j = C(f_n)$
- If $i_j = i^*$ then $C_j = C(f_y, i_j) = C(f_n)$

In other words, all individuals with incomes greater than i^* will choose to comply with the fine and individuals with income less than i^* will choose to face the consequences of nonpayment. Individuals with income of exactly i^* will be indifferent. We can then define i^* as

$$C(f_y, i_j) = C(f_n)$$

which after plugging in and solving for i, we find:

$$i^* = 1 - \frac{f_n}{f_y} \tag{2.2}$$

Because I assumed the population's income was uniformly distributed from 0 to 1, we can further define the share of the population who will choose to pay the fine as:

$$F = 1 - i^*$$

or

$$F = \frac{f_n}{f_y} \tag{2.3}$$

Intuitively, the larger the ratio between costs of the non-compliance and the costs of the fine, the more people who will choose to comply with the fine. I now consider the effects of an increase in f_n, the consequence for non-compliance on fine compliance and overall costs. Taking the derivative of the proportion of population paying the fine with respect to f_n, I first find that:

$$\frac{\partial F}{\partial f_n} = \frac{1}{f_y} \tag{2.4}$$

As we increase the consequences for nonpayment, the share of people who choose to pay the fine will increase and the rate at which this occurs will be decreasing in the fine level. Turning to costs, I find:

$$\frac{\partial C(f_j)}{\partial f_n} = \begin{cases} 1 & \text{if } \ i_j < i^* \\ 0 & \text{if } \ i_j \geq i^* \end{cases} \tag{2.5}$$

or an increase in the consequences of nonpayment increases costs among those who previously chose not to pay the fine, but creates no change among those paying the fines. Put together, we can see that increasing the consequence for nonpayment has two distinct effects. For inframarginal individuals who would previously have not paid the fine, expected costs increase at the level of f_n. For individuals whose incomes are infintesimally below i^*, the increased consequences of fine nonpayment induce fine compliance. There is no effect on the expected costs of higher-income individuals already paying the fine.

If we instead increase f_y, the fine level, the effects are nearly identical, but reversed. Specifically, we have:

$$\frac{\partial F}{\partial f_y} = \frac{-f_n}{(f_y)^2} \tag{2.6}$$

$$\frac{\partial C(f_j)}{\partial f_y} = \begin{cases} 0 & \text{if } \ i_j < i^* \\ 1 - i_j & \text{if } \ i_j \geq i^* \end{cases} \tag{2.7}$$

Increasing f_y decreases the proportion of individuals who would choose to pay the fine and this effect is increasing in f_n and decreasing in f_y. Additionally, from a cost perspective, increasing the fine level has no effect on individuals with incomes less than i^* because these individuals had already chosen not to pay the fine. It increases the costs of infra-marginal individuals who would choose to pay the fine even at the higher fine level and these marginal costs are decreasing in income.

Figure 2.3 shows these effects graphically for the case of an increase in the consequences for noncompliance.[16] As we saw above, shifting up the costs of nonpayment has two primary effects on costs. Individuals with low-incomes see overall costs increase as paying the fine remains too costly, but the cost of nonpayment has gone up. Individuals whose income is above the new i^*, but below the old i^* are now induced to pay the fine, which is more costly than not complying under the old regime, but is less costly then not complying under the new regime. There is no change for individuals whose income is above i^*. In short, increasing the consequences for non-compliance greatly increases costs for low-income individuals, moderately increases costs for middle income individuals and has no effect on high income individuals. It also leads to increased compliance with the fine, but this increased compliance is concentrated entirely among middle-income individuals.

I now turn to examining how changes in the fine level and the consequences for nonpayment affect deterrence. Let b_j denote the private benefit to an individual for committing a crime and assume it is also distributed uniformly between 0 and 1 across the population. Since costs are also constrained to be between 0 and 1 by the assumptions above, the proportion of people who will commit the crime at each income level must be $1 - C(f_j, i_j)$. Thus, we can define the total proportion of individuals, P, who will engage in an offense as:

$$P = \int_0^1 1 - C(f_j, i_j)\, di$$

[16]As shown above, effects are nearly identical but opposite for increases in the fine level.

Plugging in the cost function from above, we have:

$$P = \int_0^{i^*} 1 - f_n \, di + \int_{i^*}^{1} 1 - (f_y(1 - i_j)) \, di$$

Integrating and plugging in $i^* = (1 - f_n/f_y)$, we find

$$P = 1 - f_n + .5((f_n)^2/f_y)$$

Thus, the marginal change in crime for changes in f_n and f_y are equal to:

$$\frac{\partial P}{\partial f_n} = \frac{f_n}{f_y} - 1 \, di$$

$$\frac{\partial P}{\partial f_y} = -(\frac{f_n}{f_y})^2 \, di$$

As expected, both increases in f_n and f_y change criminal behavior. The magnitude of this change is driven by the relative levels of f_n and f_y. A higher starting ratio of f_n to f_y will lead to larger marginal changes in crime rates for increases in f_y and vice-versa. The reason is simple—when f_n is high relative to f_y most people are already choosing to comply with the fine and so the change in sanction for noncompliance affects the costs of committing an offense for relatively few individuals. When the ratio is low, the opposite is true—changing the costs of noncompliance changes the effective punishment for much of the population.

Finally, these results suggest that both increasing the fine level and increasing the consequences for nonpayment will decrease overall crime rates, but do so at the expense of different segments of the population. When the consequences for nonpayment are increased, only individuals with income less than i^* face higher costs through a combination of inframarginal individuals still committing the crime, but facing higher expected costs and compliers now choosing to not commit the crime when previously they would have enjoyed a private benefit from offending.

The opposite is true when the fine levels are changed.[17]

Although highly-stylized, this model provides important intuition for how to think about the effects of changes in fines and consequences for nonpayment on compliance, punishment and willingness to engage in criminal behavior across the income distribution. Specifically, the model makes several key predictions. First, an increase in punishment for fine nonpayment should increase average expected costs of the crime, but these costs will be distributed unevenly; only individuals with incomes less than i^* should expect to see an increase in punishment, with no change for higher income individuals. This change in punishment will take two forms. Individuals with incomes just below i^* who were close to indifferent between paying and not paying will now be induced to pay, while individuals with incomes further below i^* will now face a higher expected sanction for nonpayment. Second, the shift in costs should decrease overall unsafe driving behavior, but this reduction should be driven largely by a change of behavior among lower income individuals—individuals who would have paid the fine under the old regime see no change in expected costs and so we would expect no change in behavior. The remainder of this paper will test these theoretical predictions in the context of a large change in penalties for noncompliance with traffic offenses in Washington state.

2.5 Policy Setting

2.5.1 Background

The empirical analysis in this study uses a series of court decisions and law changes in Washington state that first invalidated and then mandated automatic license suspension when

[17]Note that if the benefits from crime reduction are independent of the income of the individual committing the crime, this implies an income-based redistribution of utility is taking place when these changes are made. If the benefits are perfectly positively correlated with the income of the offender then the redistribution only occurs between individuals with the same income but different private benefits of committing the crime. Thus, understanding the correlation between crime offenders' income and the income of those harmed by the crime is an important determinant of the equity implications of these two deterrence policies.

individuals failed to comply with the punishment for a traffic offense. Since at least the 1970s, noncompliance with the terms of a traffic offense in Washington state had resulted in the automatic suspension of a driver's license. However, in the spring of 2002, two men sued the state of Washington arguing that "mandatory suspension of a driver's license...without granting an administrative hearing violates due process." In June 2004, the Washington State Supreme Court ruled for the plaintiffs and found the existing license suspension infrastructure unconstitutional on due process grounds (*City of Redmond v. Moore* 2004). This ruling effectively invalidated all existing suspensions due to nonpayment or failure to comply with a citation as well as prevented any new suspensions from being put in place. Suspensions that occurred as a result of a direct criminal penalty rather then in response to the "failure to comply with the terms of a notice of traffic infraction, criminal complaint or citation for a moving violation" remained in effect.[18] During this period, the penalty for nonpayment of fines was ambiguous, but appears to be relegated to the threat of collection calls by the state (and eventually) debt collection agencies.[19]

This court decision was well-publicized. Numerous articles were written about the decision in all of the state's major newspapers along with op-eds and letters-to-the-editors. Patrol officers were instructed to no longer issue citations or arrest warrants for suspensions arising out of noncompliance and individuals in jail awaiting trial for DWLS charges were released. Thus, in addition to the media attention, it is likely that news of the court decision quickly disseminated through word-of-mouth as friends, neighbors and family members with suspended licenses were no longer sanctioned after being pulled over.

In 2005, the state legislature pursued a successful statutory fix to the Supreme Court ruling and mandated that all violations that took place after July 1, 2005 would again be subject to automatic license suspension for noncompliance. Unlike the initial court ruling, the new

[18]In other words, if a drunk driving conviction was punished with a 60-day suspension and fines, the 60-day suspension would remain in place, but any longer-term suspension arising from an individual's failure to pay fines would be invalidated.

[19]Further, individuals with previous suspensions who did not get their licenses reinstated would still potentially be charged with lesser offenses such as operating a vehicle without a valid license, but would not be eligible for being charged with the much more serious Driving with License Suspended (DWLS).

legislative fix received almost no attention. Newspaper archive searches reveal only a single article written about the new law in a minor local paper and there are no trend breaks in Google search activity for “suspended license” or “driver’s license.” Further, because any individual affected by the law would not be aware of the new suspension sanction for several months after their offense, it is unlikely that information could have spread quickly through word-of-mouth. Importantly, I will further show in the next subsection that there were no observable changes in the level or composition of cases before or after the law went into effect.

Under the new law, failure to comply with the terms of a traffic infraction or criminal charge would result in an automatic license suspension. Actions that trigger a “failure to comply” include ignoring a summons in response to a charge, missing hearings or noncompliance with some aspect of a penalty (typically fine nonpayment). License suspension notices are mailed to an offender’s address when a suspension goes into effect, beginning two to three months after the violation date for traffic infractions and over a longer time period for traffic misdemeanors, which may have multiple court dates prior to sentencing. License suspensions are dropped when the activity triggering the failure to comply is resolved and reinstated after a reinstatement fee is paid. In practice, suspensions are often dropped ten years after an offense even if no resolution occurs.

Figure B.1 shows total monthly driving with a suspended license (DWLS) charges before and after the policy change as a proxy for the underlying level of license suspensions in the population; it is clear that both the initial court decision and subsequent law change had a dramatic impact on license suspensions. After the court decision there was an immediate drop in the number of DWLS charges because the court invalidated all previous suspensions in addition to preventing new suspensions, depleting the stock of individuals with suspended licenses. Conversely, the 2005 law change lead to a steady increase in suspensions over an extended period of time as the stock of individuals with license suspensions gradually increased. Importantly, the first jump in DWLS cases occurs not at the July 1st threshold, but three months later when the first suspensions caused by the law went into effect. Figure B.2 shows a timeline of the major policy changes.

2.5.2 Data

The data for this analysis come from three main sources. First, I have data on all available misdemeanor and felony court cases between January 2003 and August 2018 from the Washington Administrative Office of the Courts (AOC). These data include information on all charge types, charge violation dates, filed dates, dispositions, disposition dates, sentences, fines ordered, jail time ordered, other punishments ordered, fine repayment and Failure to Appear (FTA) announcements. FTAs in the Washington court system refer to notices issued when an individual fails to comply with any aspect of his or her violation, not just failure to appear in court.[20]

The AOC also provides information on infraction cases. Traffic infractions include violations like speeding or improper passing. However, unlike the misdemeanor and felony samples, the AOC only maintains infraction records for five years following case closure (typically after all payment is received and active FTAs are closed). As a result, the infraction records provided are only those cases that were not closed prior to August 2013. Fortunately, the Washington AOC has archived monthly reports of the number of cases filed by category by month. Although these are only snapshots in time and cases can in theory be reclassified, they nonetheless provide a reasonably accurate estimate of the universe of total cases each month. Using these two data sources, I can then find the proportion of infraction cases from each month that have an outstanding FTA and outstanding suspension ten years after the violation date by combining counts of cases in my data with total case counts from the archive. This allows me to estimate the effect of the policy on punishment compliance and suspension rates of infraction offenders under the assumption that case disposal rules were correctly followed.[21]

[20]Data are provided at the charge-level for some variables (i.e. charge description and disposition), but the case-level for other variables. Accordingly, all analyses are performed at the case-level. For the vast majority of cases, the violation dates of all charges are identical. For the small minority of cases where different violation dates exist, the most recent violation date is used.

[21]Unlike for the traffic misdemeanor sample, I cannot examine the effect of the policy on individual-level recidivism outcomes because I do not observe the recidivism outcomes of individuals who paid their fines, only those whose cases remained open.

As a result, my primary analysis uses only traffic misdemeanor cases, but I also show results for traffic infractions when data are available (punishment and compliance outcomes). My primary analysis also excludes cases that consist only of license-related traffic misdemeanors. The impact of a sanction of license suspension is expected to be minimal for individuals who already have their licenses suspended (as would be suggested by being charged with a license-related offense). Additionally, there is a large increase in these cases beginning several months after the policy goes into effect as the first new automatic license suspensions arising from the policy occur. This increase does not occur discontinuously at the policy threshold, but it does imply that if these cases are included, some large differences in case composition emerge as we move several months away from the threshold (see Figure B.1). Results are generally robust to the inclusion of all cases, but given these concerns, I exclude cases consisting of only license misdemeanor charges from my primary analysis.

Second, I have data from the Washington Department of Licensing (DoL) on license suspensions beginning in 2004. These data provide information on all license suspensions in the state of Washington including individual birth-date, zip code and reason for suspension but also come with several caveats. First, for suspensions related to FTA, I do not observe if a suspension was resolved, only its issue date. Thus, I can see if a given offense ever received a suspension, but I cannot see that suspension's duration. Second, the matching between license suspensions and court records is incomplete. I restrict my suspension sample to only suspensions for FTAs, all of which contain the citation number of the offense associated with the suspension. However, in the AOC data it is clear that for data in earlier years, citation numbers repeat across some jurisdictions, but the DoL data lacks jurisdictional information for many suspensions which would allow for a unique match. When restricting my sample to two years before and after the law change, I see that 22% of citation numbers in the sample are duplicates. Accordingly, in my primary analysis, I drop all duplicate citation numbers and only consider suspensions that start within two years of the original offense (ensuring that I am not picking up duplicates that occur outside of my sample). For bias to be introduced in this setting, it would be necessary that the rate of duplicates shifts discontinuously at the threshold. In Figure B.3, I show that this is not the case.

As a separate test of the validity of this matching strategy, I also examine whether actors defined to be the same person in the license data across multiple citations are also defined as the same person in the court data. In the extreme case where the matching procedure is as good as random, the chance that different citations assigned to the same person in the license data would also be assigned to the same person in the court data would be very small. Instead, I find that 84% of multiple citations assigned to the same person in the license data are also assigned to the same person in the court data suggesting a high match rate.

Finally, I have data from the Washington Department of Transportation and California Department of Transportation on all motor vehicle crashes in both states between 2003 and 2018. These data include date, time, location of crash, driver license status (Washington only), driver insurance status, driver intoxication status, reason for crash, number and severity of injuries, vehicle make and vehicle model (Washington only). Zip-code and municipal demographic data come from the 2000 Census and 2009-2013 ACS.

2.6 Effects of Automatic License Suspensions on Compliance and Punishment Outcomes

2.6.1 Empirical Strategy

The goal of this section is to estimate the effects of changes in consequences of nonpayment of traffic-related fines on various compliance and punishment outcomes. In an ideal world, I would randomize whether or not cases were eligible for license suspensions for noncompliance and examine differences between these two groups. Of course, such an experiment is both impractical and illegal. Instead, I use the strict cut-off imposed by Washington state's legislation reinstating automatic license suspensions to approximate such an experiment with a regression discontinuity design. Individuals with violations just after the law's July 1st, 2005 cut-off had their licenses suspended for noncompliance, while no such sanction existed for cases just before the cut-off. Thus, if we assume the cut-off date is orthogonal to a defendant's willingness to comply or expected punishment, examining differences in compliance and punishment between cases just before and after the cut-off date should provide a valid estimate of the effects of automatic license suspensions.

Accordingly, I estimate:

$$Y_c = f(Time_c) + \beta Post_c + \varepsilon_c \tag{2.8}$$

where Y_c is a case-level outcome for case c, $f(Time_c)$ is a function transforming the variable $Time_c$, which measures the difference between the violation date and July 1, 2005 and ε_c is a mean zero error term. In my primary specification, I use a local quadratic regression to account for any non-linear trends induced by seasonality, but also test robustness to local linear regressions as suggested by Gelman and Imbens (2018). I use a bandwidth of one year, but test robustness to both the Calonico et al. (2017) MSE-optimal bandwidth formula (between 300 and 400 days for most specifications) and bandwidths of 3 months. Additionally, every estimate is also displayed as a regression discontinuity plot so that the reader can see for her or himself that the obtained estimate is not an artifact of specification choices. For traffic misdemeanors, the running variable ($Time_c$) is days since July 1, 2005, while for traffic infractions the running variable is months since July 2005 because these data are only available at the monthly level. All estimates are performed at the case-level. Because the same individual may be involved in multiple cases over time, standard errors are all clustered at the defendant level. Continuous outcomes such as fine levels and sentence days have several clear outliers that are implausibly large.[22] In these

[22]For fines, many of these likely come from restitution payments—for instance a greater than $50,000 fine for a traffic offense

cases, I drop all values greater than the 99th percentile of non-zero values, however results are generally robust to their inclusion.

I use two strategies to estimate the heterogeneous effects of the policy based on income, The AOC provides zip codes for each offender, but this is the most recent zip code on file not the zip code at the time of the offense. Thus, while zip code provides the most granular information about an individual's expected income level there is a concern that it may also be endogenous; for instance, exposure to the policy may lead to adverse outcomes, which may drive individuals to move to lower-income zip-codes. If these individuals are then charged with future crimes, only their new zip code will appear in my data. Although I show empirically that there is no evidence for statistically or economically significant differences in zip code median income at the discontinuity, to allay any remaining endogeneity concerns I also use an alternative method based entirely on pre-arrest characteristics.

Specifically, instead of zip code median income, I use the median income of the offender's arresting agency (i.e. the city of Bellingham if the arresting agency was the Bellingham Police Department.) I exclude cases initiated by the Washington State Patrol, which are responsible for the whole state and sheriff departments which typically have jurisdiction over unincorporated parts of a county, but may also patrol county roads within incorporated cities. Ultimately, roughly one-third of my sample is arrested by a municipal police force and is included in this analysis. The correlation between zip code median income rate and arresting agency median income is .425, suggesting that both measures are capturing similar, but not identical information.

In both cases, I divide the sample into five quintiles based on zip-code or municipal median income. Results are also robust to using poverty rates. In my main specification, I prefer the zip code specification despite the endogeneity concerns because I believe it still provides more accurate information about an offender's likely income; offenders arrested by a municipal agency need not live there and so the correlation between an offenders' true income and the median income of the municipality in which he is arrested is likely much weaker than the relationship with zip code. Nevertheless, I show that both analyses lead to substantively similar conclusions suggesting that the choice of measure is not driving the results.

2.6.2 Validity of Regression Discontinuity Design

For a regression discontinuity design to be valid, it is necessary that there be no manipulation of the outcome variable at the threshold (Lee and Lemieux, 2010). In my setting, this primarily implies two conditions: i) there was no change in behavior by either individuals or law enforcement agencies in response to the new policy that is correlated with the outcomes of interest and ii) no other changes occurred on July 1, 2005 that would have affected violation frequency or composition. It is unlikely that defendants changed behavior on July 1, 2005 because the law change received very little publicity—newspaper archive searches return only a single article that references the law's passage and it was published in the *King County Journal*, a minor paper with limited circulation. There also do not appear to be other contemporaneous law changes that may have changed compliance behavior; a search of the Washington State Session Laws in 2005 does not return any evidence for law changes that would affect punishments or definitions of traffic infractions or misdemeanors.

These qualitative explanations for lack of manipulation are supported by the empirical evidence. Figure B.4 shows the number of non-license-related traffic misdemeanor and infraction cases over time; there appear to be no discontinuous changes in number of cases filed at the July 1, 2005 threshold. If either law enforcement agencies or individuals reacted immediately to the law, we would expect to see some changes on this dimension. Figure 2.4 shows the effect of the discontinuity on the composition of traffic misdemeanor offenders and the sentences meted out for the various offenses. If the composition of offenders changed, we would expect to see a discontinuous change in offenders' criminal history. If the composition of offenses changed or judges endogenously responded to the law change by changing their punishments, we would expect to see changes in sentenced fines or jail time at the discontinuity. Again, there are no apparent differences around the threshold. Of particular note, the bottom row shows changes in zip code median income—there appears to be no significant changes at the discontinuity implying that the income-level of the reported zip code in the data was not affected by treatment. Tables 2.1 and 2.2 show these analyses in regression form; point estimates are small and statistically insignificant, again bolstering confidence that results are not driven by compositional changes or other forms of manipulation.

A final important check of the validity of the design is to examine the effects of placebo "law change" dates on our primary outcomes. Table B.1 shows the effect of a placebo "law change" on July 1st, 2006 or July 1st, 2007 on my primary outcomes of interest (FTAs, fines paid, suspensions). In all cases effects are small and statistically insignificant suggesting that there is nothing inherent about a cutoff on July 1st that creates large discontinuities in the studied outcomes.

2.6.3 Results

Theory suggests that increasing the consequences for nonpayment of fines should increase the costs of committing future crimes in two ways: a) increasing payment of fines for individuals who were previously close to indifferent between payment and nonpayment and b) increasing costs of nonpayment for inframarginal individuals. Increases in fine payment should be driven by individuals in the middle of the income distribution and increased costs of nonpayment should be driven by individuals at the bottom of the income distribution. Theory would predict little change for individuals at the top of the income distribution.

I test these predictions empirically here. I begin by examining the effects of the new license suspension regime on an offender's compliance with the terms of his or her punishment. Compliance is proxied by the presence of a Failure to Appear (FTA) notice, which is issued after an individual has failed to comply with some aspect of his or her punishment—typically nonpayment, but also ignoring hearings or summons. Figure 2.5 and Table 2.3 show the primary results. As expected, the threat of license suspension leads to a large and persistent decrease in outstanding FTAs. The effect grows from a 3 percentage point decrease (10%) after 2 years to 5 percentage point decrease (22%) after ten years. These effects are all highly statistically significant. The bottom right panel of Figure 2.5 show a similar effect for traffic infractions; after the license suspension regime begins, open FTAs at ten years post-violation fall by about 6 percentage points (40%). These effects are all robust to different specification, weighting or bandwidth choices (Table B.2).

The adoption of the license suspension regime causes a meaningful proportion of offenders to increase compliance with the terms of their offense. The theory developed above predicts that this increased compliance will occur in a predictable manner; very rich and very poor offenders should exhibit little change in compliance behavior because most rich offenders will already be complying and poor offenders will still find it too costly to comply. Conversely, middle income offenders should be induced to begin compliance because they were likely to be the marginal noncompliers prior to the policy change. I test that theory here by examining differential effects by zip-code median income. Note that because all zip codes contain some mix of low, medium and high-income individuals, the effects by zip code should be much noisier than what is predicted by the theory. Nevertheless, the results shown in the left panel of Figure 2.6 are broadly consistent with the theory's predictions; offenders from zip codes with low or high median incomes experience relatively little change in behavior. Conversely individuals from zip codes near the middle of the income distribution see a large increase in compliance behavior. The right panel of Figure 2.6 shows that we see a similar pattern when performing the same analysis using municipal law-enforcement agency median income—again the largest reduction in noncompliance comes from the center of the income distribution.

Figure 2.7 and Table 2.4 show the effects of the license suspension policy on the actual amount of money paid for each case. Consistent with the compliance results above, the license suspension policy significantly increases the amount of money paid for a traffic misdemeanor offense as well as the probability that any amount is paid.[23] Specifically, the threat of license suspension increases the total amount of money paid for an offense by approximately $67 (10%) and the probability of any repayment by 3 percentage points off a base of 76 percent. Further, as a placebo test Columns (4) and (8) show the effects of the policy on fine repayment for non-traffic criminal misdemeanors, which were never subject to license suspension for nonpayment. Reassuringly, we see that the policy has no effect on this group, again providing us with increased confidence that the observed effects are indeed causal.[24]

I next estimate the effect of the license suspension policy on the costs of fine noncompliance for inframarginal individuals, which should occur largely through increased license suspensions. Figure 2.8 shows the main results. The left panel of Figure 2.8 demonstrates that unsurprisingly, the license suspension policy led to a large increase in the number of traffic misdemeanor cases that had a suspension by two years after an offense.[25] Although this effect is somewhat mechanical as suspensions were banned for offenses that occurred before the law change, the magnitude of the effect is important; for traffic misdemeanor offenders there was a greater than 20 percentage point increase in the probability that a suspension would ever be attached to their case in the new policy regime. Among traffic infraction offenders (right panel) the effect is similar, but much smaller (6 pp increase), reflecting the fact that baseline fine compliance rates are much higher for infractions that misdemeanors. Unfortunately, I lack data to see how long these suspensions persist. However, given that nearly 20% of individuals with offenses after July 1, 2005 still had active FTAs ten years after the offense, it seems likely that the vast majority of these suspensions also remained active for multiple years.

The left panel of Figure 2.9 shows how this effect differs by zip code median income for traffic misdemeanor offenders. Consistent with the theory developed above, the license suspension

[23]Data limitations prevent examining the effect on traffic infraction offenses.

[24]I cannot do parallel placebo checks for FTAs because data on FTAs for non-traffic misdemeanors are not available. I cannot do a parallel analysis for suspensions because noncompliance with non-traffic misdemeanors does not lead to license suspensions in Washington state.

[25]Note that this is the proportion with a suspension tied to their specific case, not the probability that the offender has a suspension due to any offense.

policy leads to a much larger increase in license suspensions among offenders from lower-income zip codes—the policy causes suspensions due to FTA for offenders in these zip codes at nearly double the rate of those from the highest-income zip codes and this effect decreases linearly in zip code median income. The right panel of Figure 2.9 shows that results are similar if we instead use law-enforcement agency service area median income.

Finally, I examine the effects of the policy on the major legal downstream consequences of license suspensions: driving with license suspended (DWLS) charges. As discussed above, DWLS are class A misdemeanors and can lead to jail time. Unlike with the outcomes above, future DWLS offenses are an individual, not a case-level outcome. As a result, a regression discontinuity estimate of these effects based on whether or not a violation occurred before or after the new law only estimates the effect of having an additional case under the new law, not the effect of being in a new punishment regime going forward.[26] Individuals with cases before the threshold may (and often do) quickly accumulate cases after the threshold, which expose them to the new punishment regime, "diluting" the effect of treatment. Accordingly, I instead instead estimate the effect of the new law's legal consequences using time-series and difference-in-differences designs.

Figure B.5 first provides some descriptive evidence of the magnitude of the new law's effect by comparing average total charges, fines and sentenced days for license misdemeanors and non-traffic misdemeanors such as shoplifting and disorderly conduct in each court before and after the law comes into effect. All three outcomes for license misdemeanors increase dramatically after the new law relative to non-traffic misdemeanors, with a particularly large effect for charges and fines. Prior to the law coming into effect, courts on average had four times as many non-traffic misdemeanor charges as license misdemeanor charges and collected twice as many dollars in fines. By two years after the law, courts had only 1.3 times more non-traffic misdemeanor charges and collected similar amounts in fines from both case types.

[26]For instance, an individual with a speeding event before the threshold and after the threshold face the same potential penalty for any new traffic offenses that occur after the threshold.

While suggestive, this analysis only shows trends over time across crime types—it is in theory possible that other contemporaneous events could be driving these results. To address this concern, I next perform a separate analysis comparing total charges, total fines, sentenced jail days and served jail days for license misdemeanors by zip code median income under the assumption that lower income individuals would be much more affected by changes in consequences to noncompliance than higher income individuals.[27] Figure 2.10 shows the main results. For context, the difference between the 10th and 90th percentile of median income is roughly $30,000. Thus, relative to a zip code in the 90th percentile of income, two years following the law change, a zip code at the 10th percentile would have 7 additional license misdemeanor charges/1,000 population per year, an additional $2,400 per 1,000 population per year in license misdemeanor fines, an additional 480 sentenced days in jail per 1,000 population per year and an additional 50 served days in jail per 1,000 population per year.[28]

Figure B.6 shows that results are qualitatively similar when using municipal median income suggesting that results are not driven by endogenous sorting due to the policy. Figure B.7 shows that this increase in fines and sentenced days is not simply crowding out penalties from other types of crime—when we look at the effect on total misdemeanor fines and sentenced days, results persist. Together, these results suggest that the license suspensions induced by the new policy had large disproportionate effects on low-income residents.

Combined with the compliance results, these estimates suggest that imposing a license suspension penalty for non-compliance with traffic offenses imposes large costs and that these costs are highly heterogeneous by income. Low-income individuals exhibit smaller changes in their compliance behavior, but see a large increase in the severity of their total punishment

[27] I restrict my sample to only areas with greater than 5,000 population in order to get consistent quarterly estimates. Results are robust to including all zip codes

[28] Most sentenced jail days for DWLS (and other traffic misdemeanors) are suspended. If an offender completes a probationary period without any violations, these days will never be served. However, if the offender violates their probation, even for a minor offense, the judge can order the offender to serve some or all of this suspended sentence. Even in the absence of serving the sentence, the threat of suspension is a potent cudgel that gives judges significant power over defendants' lives following the sentence and so sentenced jail days should be thought of as costly in and of themselves.

as shown through increases in license suspensions and driving with a suspended license cases. Medium-income individuals are more likely to be induced to comply with the policy, but also see some increase in their effective punishment. Finally, high-income offenders face a much smaller change in costs.

It is lastly important to note that a large potential unmeasured cost of the policy change is the effect of license suspensions on employment or other measures of financial stress. Unfortunately, given data limitations it was not possible to estimate the effects on these outcomes using the regression discontinuity framework here. However, recent work by Mello (2018) finds that short-run (30 day) license suspensions for unsafe driving leads to increased financial distress and decreased levels of employment. Mello (2018) cannot causally identify the effect of automatic suspensions for nonpayment studied here, but given the extreme length of these suspensions (many last up to ten years), it is likely their effect on financial stress is significantly higher. Estimating these costs is an important area of future research.

2.7 Effects of Automatic Suspensions on Traffic Safety

I next examine the effect of the license suspension policy on changes in driving behavior. The previous section showed that the license suspension policy increased the costs of engaging in illegal traffic behavior and that these increased costs were larger for lower-income individuals. Theory predicts that these increased costs should decrease unsafe driving. These decreases could occur either through individuals without suspensions driving more safely because they are aware that a consequence of a future driving violation may now be a license suspension or through an increase in the number of individuals with license suspensions, who presumably will reduce unsafe driving because the consequences of driving at all (and particularly unsafe driving) has increased. In this section, I first focus on the total effect of automatic license suspensions on driving behavior because this is the relevant parameter from a policy perspective. I then attempt to perform some suggestive tests to shed light on the underlying mechanism driving the results.

2.7.1 Empirical Strategy

In this section, my objective is to estimate the extent to which changes in the consequences for noncompliance with traffic punishments affects overall levels of road safety. To estimate this object causally, I would ideally like to have some individuals randomly assigned to have their licenses suspended if they do not comply with the terms of a traffic offense and other individuals randomly assigned to not be eligible for license suspension. Note that this is not what the regression discontinuity design used in the above section does; it instead approximates a randomization of different *cases* to different policies, but not individuals. All individuals regardless of whether their offense was before or after the new law came into effect are exposed to the new punishment regime for any future offenses that occur after July 1, 2005. Thus, while the regression-discontinuity design is ideal for estimating the effect of the policy on case-level outcomes such as fine-repayment or license suspensions, it will not provide an estimate of the average deterrence effect of the policy.[29] Instead, to estimate the effect of the law on future driving violations, I will use a series of difference-in-differences and triple-difference models aimed at comparing the driving behavior of groups expected to be differentially affected by the changes in the consequences of noncompliance before and after the policy change.

I begin by analyzing the effects of the law changes on driver safety by investigating changes in injury-producing crashes during the period in which license suspensions for noncompliance were unconstitutional relative to both periods before the court decision and after the law

[29]Instead, the regression discontinuity design recovers the effect of having one offense after the law change. Having a post law-change offense leads to an increased likelihood of license suspension and may also lead to increased awareness of the new policy. Thus, if much of the effect of the new policy is driven by license suspensions, we may still uncover an effect using the regression discontinuity design. Conversely, if much of the effect of the policy is driven by a general change in driving behavior by individuals who are at risk of suspensions then we would likely observe no effect in the regression discontinuity design despite the policy having an overall deterrence effect. In either case, any effect of a post-policy offense on recidivism will likely decay as “control” individuals with offenses in the pre-policy period begin accumulating offenses after the new law goes into effect, thereby “diluting” any treatment effect of an additional post-law change offense.

change for low relative to high-income drivers.I use only crashes that result in injuries to address endogenous reporting concerns. I proxy for driver income using vehicle age under the assumption that lower-income drivers drive older vehicles—I also show robustness to using imputed vehicle value. The idea underlying this analysis is that if vehicle age is correlated with driver income and if increasing the sanction for noncompliance disproportionately affected low-income drivers, then the average age of vehicles in crashes when automatic suspensions are eliminated should increase if drivers are indeed responding to this decreased punishment. To control for other correlated trends that may be separately affecting low-income drivers during this period, I use vehicles involved in crashes in Northern California, a region that is similar demographically, economically and geographically to Washington state as a control group.[30] My sample starts in 2003, the first year Washington has data available, and ends in July 2008, three years after the law reinstating license suspensions came into effect. My primary specification is:

$$Y_{dat} = \tau_a + \beta Policy_t * Washington_d + Z_{at} + \varepsilon_{dat} \tag{2.9}$$

where Y_{dat} is the age of the car of driver d who was in an accident in the jurisdiction of agency a on day t, τ_a is an agency (i.e. the law enforcement agency that responded to the crash) fixed-effect, $Policy_t$ is a vector of indicators for the policy regime a crash occurs under, $Washington_d$ is an indicator for whether or not a crash occurred in Washington state, and ε_{dat} is a mean zero error term. In the base specification, I further include agency type (state police, county, municipal) by date fixed-effects (Z_{at}) to ensure that we are only comparing changes in crash composition within the same types of places (i.e. interstates, county roads, local roads) in Northern California and Washington.[31] All vehicles older than 20 years were excluded as outliers.

[30]The state of Oregon may be a better control group, but unfortunately there is no crash-level data available for the time period studied.

[31]Generally, state police are responsible for highway crashes, sheriff departments are responsible for crashes in rural areas and county roads, and municipal police departments are responsible for crashes within incorporated areas.

The identifying assumption in this analysis is that absent the policy changes and conditional on covariates, trends in the likelihood that a crash involves differently-aged vehicle should be identical between Washington and Northern California. The most obvious potential violation to this assumption is that low-income drivers in Northern California are not a good counterfactual for drivers in Washington. I attempt to control for this potential problem by including an increasing number of fixed-effects that restrict the sample to observably-similar locations in Washington and Northern California. I begin by introducing county population decile by date and county poverty rate decile by date fixed-effects, but also show robustness tests using municipal population and poverty rate deciles when restricting the sample to only crashes that come under the jurisdiction of municipal police departments.

I also further test this assumption by performing the same estimation as above using a dynamic difference-in-difference design to see if any pre-trends exist prior to the first policy change in June 2004 and whether effects begin to converge back towards zero in the period after the July 2005 law begins to take effect. The combination of the introduction and ending of the non-suspension policy between the court-decision and the law change proves a stronger than normal test of parallel trends and I find no evidence for any violation. An additional concern is that even after restricting my sample to only injury-producing crashes, results may still be driven by endogeneity in reporting. Thus, in addition to my primary analysis using only injury-producing crashes, I separately report results for only crashes with serious injuries—results are much noisier given the smaller sample size, but qualitatively similar.

A final concern is that perhaps any observed differences are simply being driven by changes in the composition of the vehicle stock in Washington state relative to California. To address this potential issue, I perform a second test comparing changes in crashes per capita in high and low poverty jurisdictions in Washington state relative to California using a triple difference design. Reassuringly, results are highly consistent with the vehicle-level regression.

2.7.2 Results

Table 2.6 shows the primary result of the traffic accident analysis. The outcome variable is a proxy for driver income: the age of a vehicle involved in an injury-producing crash. Columns (1), (3) and (5) show the effects of being in the period during which suspensions were unconstitutional relative to both the year and a half before the court-decision and three years after the law change reinstating license suspensions. Column (1) shows the base analysis with agency type (i.e. highway patrol, sheriff or municipal police) by date fixed-effects and jurisdiction by county fixed-effects. Column (3) shows results after adding in controls for county population decile by date and county poverty decile by date fixed-effects to try to compare crashes in comparable counties in Northern California and Washington. Column (5) limits the sample to even more similar localities by comparing only crashes that occurred in the jurisdiction of municipal police forces and including municipal population decile by date and municipal poverty rate by date fixed-effects.[32] Across all specifications, we see that the average age of vehicles involved in injury-producing crashes during the period in which suspensions were unconstitutional increased by about .15 years, a roughly 2% increase. Coefficients are similar across all three specifications.

Columns (2),(4), and (6) show parallel specifications, but here I examine the effects of being in either the period prior to the court decision, the first year after the law change, the second year after the law change or the third year after the law change relative to the period in which suspensions were invalidated. If the observed effects are truly driven by the license suspension policy we would expect that average age of vehicles in injury-producing crashes be lower both before the court-decision and after the new law goes into place relative to the period in which suspensions were unconstitutional. That is precisely what we see in the data—the average vehicle age in a crash is significantly lower both before the court decision and after the new law comes into place. Further, average vehicle age converges slowly back to the pre-decision mean,

[32]Because county sheriffs and state police patrol both inside and outside of incorporated areas, I exclude them from this analysis.

consistent again with the new law not being well-publicized. Figure 2.11 shows this pattern at the half-year level in event-study form. Although noisier, these results demonstrate the same general pattern; an increase immediately after the court decision and then a slow convergence back to the pre-court decision mean. Note that this is a much stronger than normal test of the parallel trends assumption as the policy turns both on and off. Thus, the fact that we observe this pattern increases confidence that the observed effect is indeed causal.

Table B.3 show these same analyses, but using estimated vehicle value rather than vehicle age as the outcome variable.[33] Results are substantively similar; the period in which license suspensions are unconstitutional leads to a $150-$300 decrease in the value of vehicles involved in crashes or a 1%-2% decline. Table B.4 shows the effects of using different accident-severity cut-offs. Results using only fatal or serious accidents where any reporting decision should be minimal are larger, but much noisier, consistent with the much smaller sample sizes.[34] The effect when using a sample of accidents that include even the possibility of an injury is smaller (.8% decline), but remains statistically significant. In general, effects appear larger the more restrictive the criteria, which provides suggestive evidence that these results are not driven by endogenous reporting of accidents. Finally, Table B.5 shows the effects when using all of California as a control group rather than just Northern California. Results are broadly similar, especially in the specifications adjusting for covariates, suggesting that effects are not driven by changes in Northern California alone.

In addition to looking at the average age of a vehicle in a crash, we can also look at how crash rates varied based on the poverty level of a given jurisdiction. Theory would predict that we should see higher crash rates in higher poverty jurisdictions in Washington state during the period in which the Supreme Court decision was in effect. Accordingly, I perform a triple

[33]We use vehicle age as our primary outcome variable because the vehicle value estimation process requires relatively strong assumptions about depreciation curves within various makes and models and insufficient information is available to provide estimation for a substantial subset of vehicles. See Appendix B.2 for more information about how these values are estimated.

[34]Given the small sample size, I do notinclude specifications with additional covariates with these outcomes.

difference design comparing municipalities above the median poverty level to those below the median poverty line in Washington state and Northern California.[35]

Table 2.7 shows the results of this analysis; being above the median poverty line in Washington state leads to a .17 per 1,000 population increase in the number of cars involved in injury-producing crashes per month when the Supreme Court decision was in effect, a roughly 12% increase. As with the previous analysis, crash rates in high-poverty jurisdiction were lower both before the court decision *and* after the new law came into effect (Column (2)), providing additional evidence that the changes were caused by the shift in suspension policy and not a broader secular trend. Results are qualitatively similar, albeit statistically insignificant when using the poverty rate of a jurisdiction instead of a binary indicator for above or below median (Columns 3-4). Particularly reassuring, results are also robust to performing a difference-in-differences within Washington state alone (Columns 5-6) suggesting that nothing about the California counterfactual is driving our results. Finally in Columns (7)-(8) we can see that results are broadly similar when using a 1,000 instead of 5,000 population threshold for inclusion into the analysis.

Table B.6 shows the results of a similar analysis, but with total traffic accident injuries per 1,000 population per month as the outcome variable. Effects are proportionally similar; the court decision led to a 9% increase in accident victims per capita in higher poverty districts, but are statistically insignificant (p=.25). Together, these results provide additional evidence that the existence of these policies can have important implications for road safety.

To conclude, I examine who reaped the benefits of these increases in traffic safety. If low-income individuals largely get in accidents with one another then the policy change would be redistributive within each income group from unsafe drivers to safe drivers, but have no cross-income distributional consequences. However, if driver incomes are uncorrelated with the

[35]The primary analysis includes only jurisdictions with population greater than 5,000 population because smaller jurisdictions have large amounts of inter-month variation in crashes per capita, which introduces a substantial amount of noise into the estimates. However, results are robust to using lower (i.e. 1,000 population) cut-offs.

income of other drivers in an accident then the observed reduction in bad driving by low income individuals is redistributive both from safe to unsafe drivers and from low-income to high-income drivers because the costs only affect unsafe, low-income drivers, but the benefits accrue to both low and high-income drivers. Figure B.8 shows that indeed there is very little correlation between vehicle age among vehicles involved in a multi-car traffic accident. Increasing the age of the at-fault vehicle from 1 to 15 years is associated with only a .2 year increase in the value of the non-at-fault vehicle. Thus, while the costs of the increased deterrence were largely felt by low-income drivers, the benefits in road safety were shared across the income distribution.

2.7.3 Mechanisms

The above deterrence effects could be caused either by low-income individuals driving less (or more safely) in response to knowledge that the consequences of a traffic violation has increased or because more low-income individuals now have license suspensions, which themselves raise the costs of driving and particularly driving unsafely. From a policy perspective, we only care about the total deterrence effects estimated above. However, from a theoretical perspective we may be interested in the extent to which these results are driven by unsuspended individuals' fear about increased future consequences for unsafe driving through increased costs of noncompliance relative to the increase in license suspensions levels. In other words, one could consider the following thought experiment: if a penalty for noncompliance existed that incurred the same utility cost as license suspensions (i.e. wage garnishment), but had no independent consequences on driving, how much of the increase in safe driving would persist?

Answering this question is challenging given the existing data. Ideally, I would estimate the effect of the court-decision invalidating suspensions on changes in driving behavior of individuals without license suspensions, but who had different underlying likelihoods of noncompliance—however because all suspension records were purged following the court decision, the data to perform this analysis does not exist. Similarly, I could estimate whether individuals with different propensities for noncompliance had different rates of accumulating their first post-law change offense—this first offense could not possibly be influenced by suspensions because the court-decision invalidated all previous suspensions. However, because I lack historical data on infraction cases, which are a large driver of suspensions, such an estimation is not possible.

Instead, I perform a more suggestive test aimed at uncovering the extent to which suspensions may limit driving (or lead to safer driving). I use more recent data where I have the full universe of traffic infractions and take advantage of the fact that noncompliance with infractions and misdemeanors only officially occurs several months after the case is filed as individuals are given a grace period to comply. Accordingly, I create a sample of all individuals with license suspensions for noncompliance between 2014 and 2016. I then examine violation behavior in the 2 months before and 4 months after the suspension is in place (2 months is the minimum time following a case filing in which a suspension would occur). I exclude cases in which the suspension occurs more than 6 months after the filing date ($\approx$ 10% of cases). If suspensions were driving the observed results, we would expect traffic violations and misdemeanors to fall sharply following the suspension. Individuals are notified of suspensions by mail and so should be generally aware that their license is now suspended, All outcome variables are adjusted for day-of-week, month and days since the case was filed to ensure I am not picking up other secular trends.

As the top and middle panels of Figure B.9 show, there is no evidence of changes in traffic offenses following a suspension—there is neither a discontinuity around the suspension day threshold, nor any major change in slope following the suspension for either traffic infractions or traffic misdemeanors. The results for traffic misdemeanors are particularly reassuring as it is likely that police officers have less discretion in citations for more serious offenses. These results suggests at least in the shorter-run, suspensions are not driving safer behavior among suspended individuals in this population. Finally, as a sanity check that suspensions are indeed taking effect, the bottom panel of Figure B.9 shows changes in annualized daily rates of license misdemeanors relative to the suspension date. Suspensions have a large effect on this outcome; relative to the days just before the suspension goes into place, individuals have a .45 charge per year increase in the probability of having a license-related misdemeanor charge. Thus, the lack of a large driving response is not due to the suspension for some reason not coming into effect. Although far from definitive, the above analysis suggests that at least in the short-run, license suspensions do not cause large improvements in driving behavior—it is not therefore implausible that a large proportion of the effects observed in Section 6.2 could be driven by changing perception of costs rather than the suspensions themselves.

2.8 Conclusion

Every year, the American criminal justice system processes tens of millions of criminal and civil cases for violations of misdemeanor criminal offenses and traffic violations. These crimes have large effects on societal health and well-being. Unsafe driving causes traffic accidents, which lead to property loss, injury and death. Minor misdemeanor offenses such as retail theft, low-level assault and criminal mischief harm local businesses, reduce property values and degrade a neighborhood's quality of life. These offenses are primarily punished by fines and other forms of non-incarcerative punishments. Because compliance with non-incarcerative punishments is an individual choice, the level of deterrence these punishments create will depend not only on the level of the court-ordered punishment, but also the costs of punishment noncompliance. These fines and associated consequences for noncompliance also impose significant costs on offenders including financial distress and loss of employment (Mello, 2018). However, despite the high levels of noncompliance in the United States and the high social costs of both these crimes and their associated punishments, little theoretical or empirical evidence exists on how changing noncompliance costs affects compliance, overall punishment and total deterrence.

This paper begins to fill this gap in knowledge. I show theoretically that increasing costs of noncompliance should deter individuals from engaging in the original criminal offense, but the costs of crime reduction will be borne disproportionately by the low-income population. I next show that these theoretical predictions are supported empirically using a natural experiment in Washington state whereby offenders with violations after a certain date were eligible for automatic driver license suspension if they did not comply with the terms of their punishments. I demonstrate that this increase in noncompliance costs led to large increases in compliance and in the effective punishment faced by traffic offenders. I further show that the increased noncompliance costs decreased unsafe driving from low-income residents as demonstrated through fewer traffic misdemeanors and injury-causing traffic accidents, while the benefits from these improvements in road safety were shared across the income distribution. This suggests that increasing costs of noncompliance through license suspensions can be an effective, but highly regressive way to improve road safety.

However, these findings have important applications beyond traffic safety. In the United States, recent research has emphasized a "criminalization of poverty", whereby the inability of low-income individuals to pay small fees or fines has dramatic consequences that have the potential to restrict individuals' future income-earning possibilities. One rationale for such "criminalization" is that without such consequences, low-income individuals face little effective punishment for many lower-level crimes punished by fines. In a sense, absent such punishments for noncompliance, these individuals would face a situation akin to the judgment-proof problem for small firms. In this paper, I show that in some ways this argument is supported empirically; increasing the costs of noncompliance does indeed increase compliance with punishments for a given law and deter lower-income individuals from engaging in unsafe driving. However, such increased noncompliance costs also lead to dramatically higher punishment costs for low-income individuals even though the benefits of the policies are widely-shared, creating important equity concerns.

One potential way to address this inequity without dramatically reducing noncompliance costs is to set fines and other non-incarcerative punishments in ways that work to equalize utility costs across the income distribution. While doing so in an exact way is impossible, adopting the "day-fine" method used in many Northern European countries, whereby fines are assessed at a certain proportion of an individual's income (Kantorowicz-Reznichenko, 2015), is a promising way forward. Alternatively, US states could improve processes for individual judges to change fine levels based on an individual's income. In most states, judges already have this prerogative, but in practice, these decisions are ad-hoc, have the potential for significant bias and are sparingly used. Systematizing this process while still maintaining significant costs for noncompliance would go a long way towards reducing the inequity of the current system without sacrificing the benefits of deterrence. In general, future research on how to better adapt punishments for low-level crimes such that they create sufficient levels of deterrence without exacerbating existing inequalities is essential to create a more effective and more fair American criminal justice system.

Chapter 2, in full, is currently under submission for publication of the material. Krumholz;Samuel. "Enforcing Compliance: The Case of Automatic License Suspensions". The dissertation author was the primary investigator and author of this material.

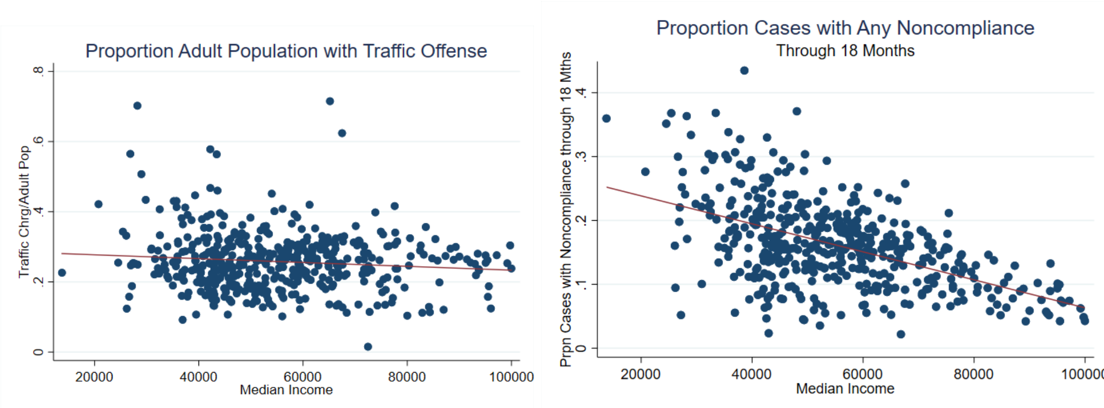


Figure 2.1: Traffic Offense and Noncompliance Rates by Zip Code Poverty Level (Offenses between 2014-2016)

This figure shows the share of 18-and-over population with any traffic offense (left panel) and the proportion of individuals with traffic offenses who were out of compliance with the punishment 18 months after the offense (right panel). Traffic offenses are all offenses coded by the Administrative Office of the Court (AOC) as traffic infractions or traffic misdemeanors. Only individuals with offenses between 2014 and 2016 are included. Zip codes outside the state of Washington, that do not match with a Zip Code Tabulation Area (ZCTA) or with fewer than 1,000 residents are excluded. All data come from the Washington Administrative Office of the Courts (AOC).

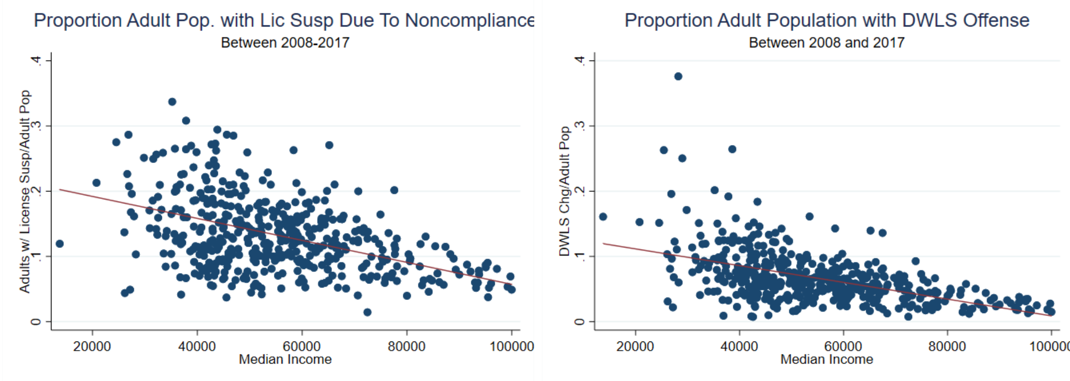


Figure 2.2: Poverty Rate, License Suspensions and Driving with License Suspended Convictions

This figure shows the share of 18-and-over population with any license suspension (left panel) and driving with license suspended charge (right panel) in 2008-2017 by zip code median income. Zip codes outside the state of Washington, that do not match with a Zip Code Tabulation Area (ZCTA) and with fewer than 1,000 residents are excluded. Data come from the Washington Administrative Office of the Courts and the Washington Department of Licensing.

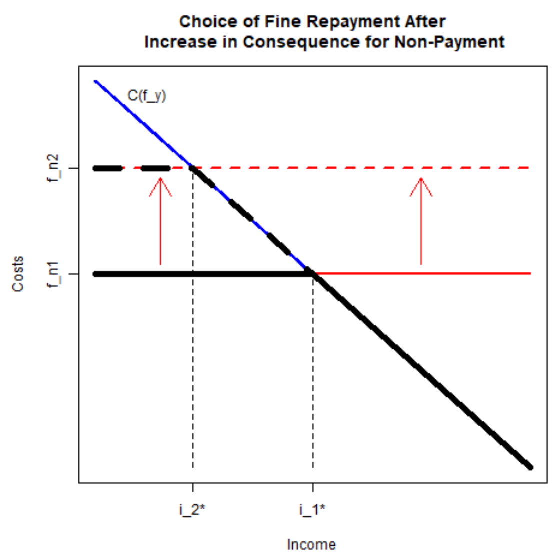


Figure 2.3: Changes in Fine Repayment Caused by Changes in Consequences for Non-Payment (f_n)

This figure shows how theoretical predictions of total punishment and fine repayment change in response to an increase in the costs of noncompliance from f_{n1} to f_{n2}. In response to the shift in noncompliance costs, all individuals with incomes less than i_{2*} experience in increase in total costs of punishment equivalent to $f_{n2} - f_{n1}$. All individuals with incomes between i_{2*} and i_{1*} are now induced to comply with the fine and experience an increase in total punishment equal to $C(f_y) - f_{n1}$. All individuals with incomes greater than i_{1*} see no changes in behavior or expected costs.

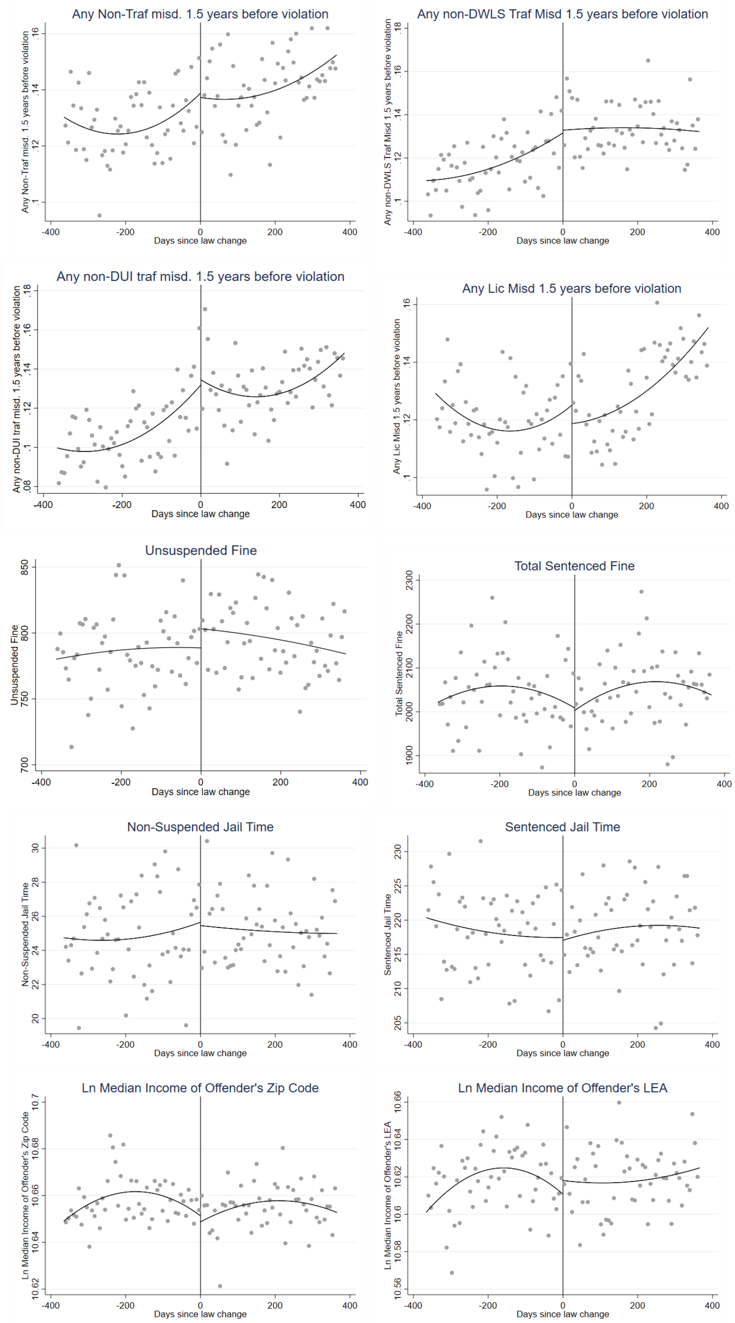


Figure 2.4: No Change in Offender Composition or Case Attributes across Threshold

Figure shows changes at the July 1, 2005 discontinuity in offender previous criminal history in the year and a half prior to the offense (the maximum length possible given the start of sample), violation fine and jail sentences, offender zip code and law enforcement agency service territory median income. All demographic data come from the US Census. All data on criminal history and sentencing come from the Washington Administrative Office of the Courts. Values greater than the 99th percentile of non-zero outcomes were excluded as outliers for all non-binary outcomes, as the distributions are heavily right-skewed. The sample consists only of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses.

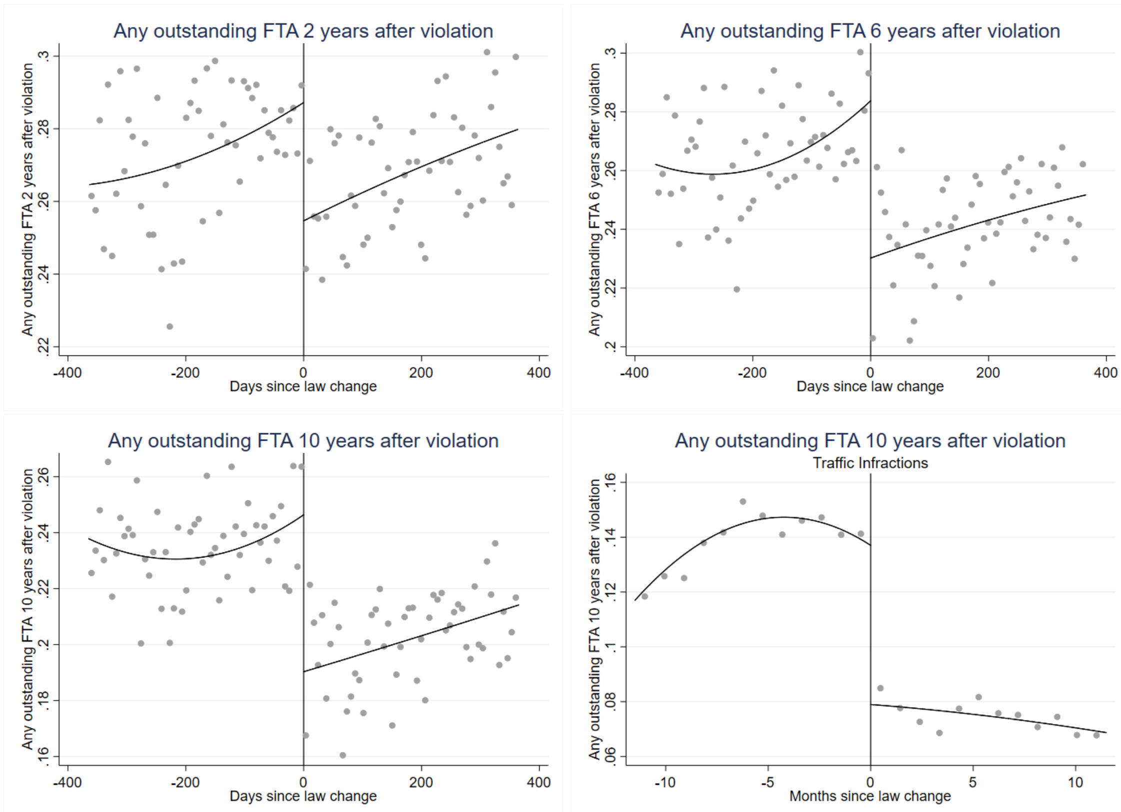


Figure 2.5: Effect of Discontinuity on Outstanding FTAs: Traffic Misdemeanors and Infractions

Figure shows the effect of the discontinuity on the existence of an outstanding FTA various time periods after the original violation. A violation is considered to have an outstanding FTA if an FTA is open at the date described (i.e if two years, then is considered to have an FTA if FTA is open exactly two years after violation). FTAs are issued for failure to appear in court, failure to respond to a citation or failure to comply with a punishment. Outcomes are binned at the weekly level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses (top panel, bottom-left panel) or any traffic infraction (bottom-right panel). All data come from the Washington Administrative Office of the Courts.

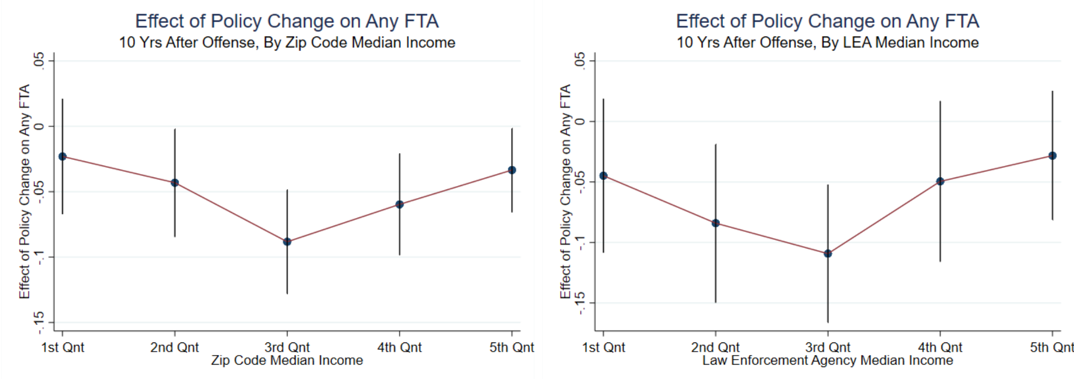


Figure 2.6: Effect of Suspension Policy on FTAs by Income

Figure shows the effect of the discontinuity on the existence of an outstanding FTA ten years after the original violation. FTAs are issued for failure to appear in court, failure to respond to a citation or failure to comply with a punishment. Coefficients come from five separate regressions for zip code (left) or law-enforcement agency service territory (right) quintiles of median income. Median income levels come from the 2000 US Census. All regression discontinuity models use a bandwidth of 1-year and local quadratic regression with a triangular weights. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All data come from the Washington Administrative Office of the Courts. Standard errors are clustered at the defendant level.

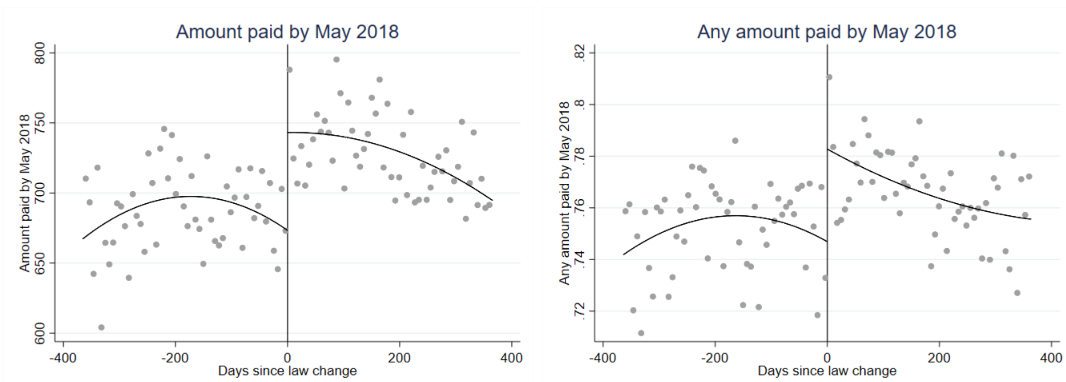


Figure 2.7: Effect of Discontinuity on Fine Repayment: Traffic Misdemeanors

Figure shows the effect of the discontinuity on the amount of money paid for an offense by ten years after the original violation. Outcomes are binned at the weekly level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses unless otherwise noted. All data come from the Washington Administrative Office of the Courts and are as of August 2018. Values greater than the 99th percentile of non-zero outcomes were excluded as outliers for all non-binary outcomes, as the distributions are heavily right-skewed.

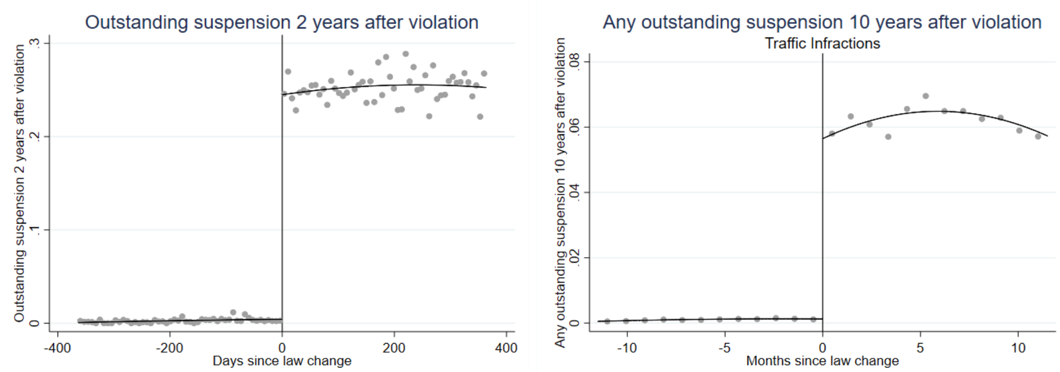


Figure 2.8: Effect of Threat of License Suspension on Actual License Suspension

Figure shows the effect of the discontinuity on whether or not a case had a license suspension attached to it for an FTA two years after the violation date. FTAs are issued for failure to appear in court, failure to respond to a citation or failure to comply with a punishment. Outcomes are binned at the weekly level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses (left panel) or any traffic infraction (right panel). All case data come from the Washington Administrative Office of the Courts and all license data come from the Washington Department of Licensing (DoL). License suspensions and case data were merged using citation numbers; approximately 22% of citation numbers in the court data were duplicates across jurisdictions and so were dropped from the sample as a unique match could not be found.

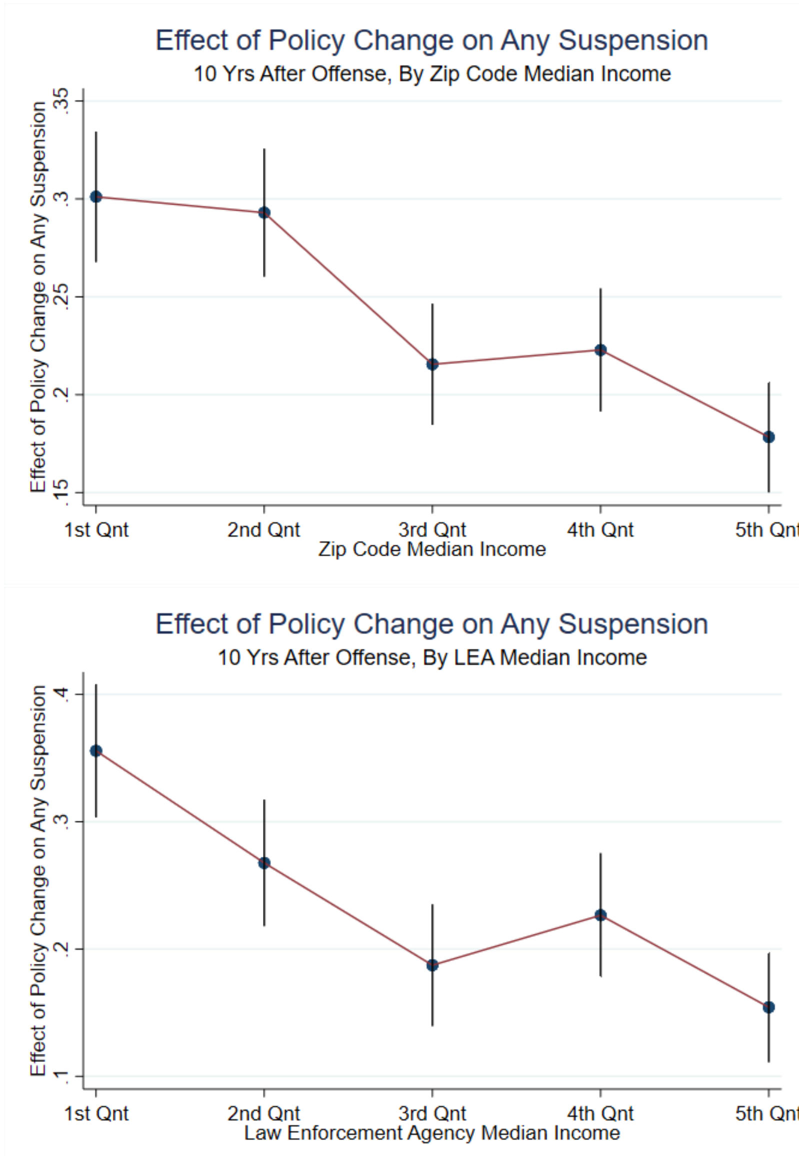


Figure 2.9: Effect on Suspensions by Income

Figure shows the effect of the discontinuity on the probability that a case has a license suspension attached to it two years after the original violation by zip code (left) or law-enforcement agency service area (right) median income. Coefficients come from five separate regressions for zip code (left) or law-enforcement agency service territory (right) quintiles of median income. Median income levels come from the 2000 US Census. All regression discontinuity models use a bandwidth of 1-year and local quadratic regression with a triangular weights. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charged unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts and all license data come from the Washington Department of Licensing (DoL). License suspensions and case data were merged using citation numbers; approximately 22% of citation numbers in the court data were duplicates across jurisdictions and so were dropped from the sample as a unique match could not be found.

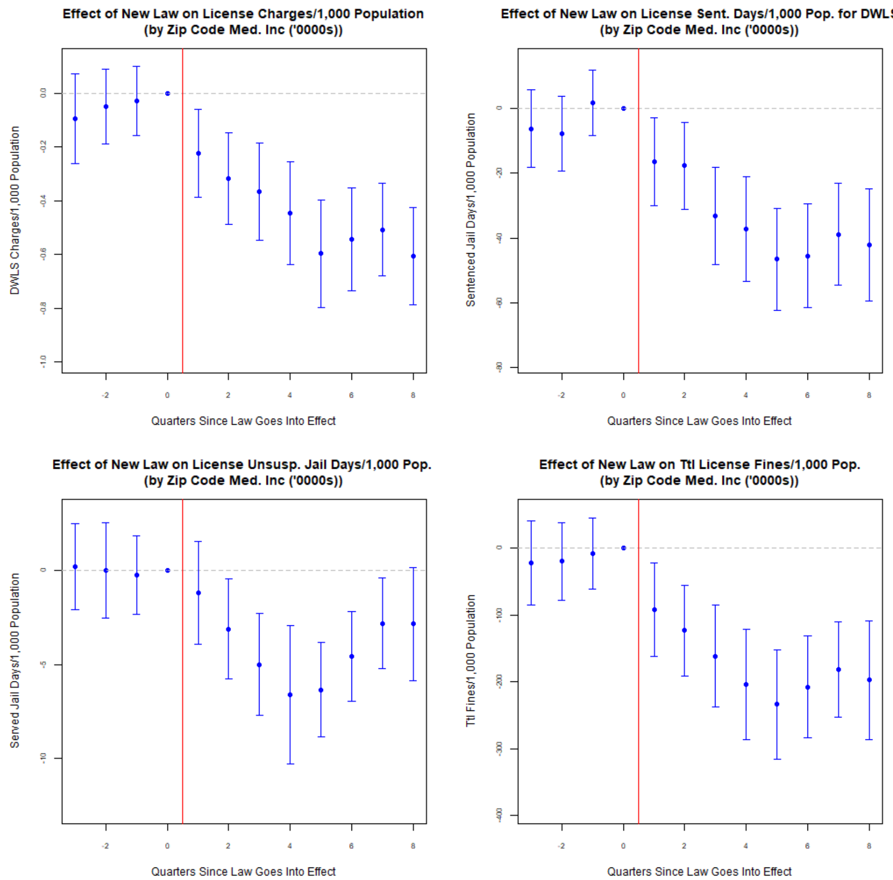


Figure 2.10: Effect of New Law on License Misdemeanor Charges, Sentenced Jail Days and Fines by Zip Code Median Income

The figure shows the coefficients from a regression of zip code monthly license charges per 1,000 population, license misdemeanor fines per 1,000 population, license misdemeanor sentenced jail days per 1,000 population and license misdemeanor served days per 1,000 population on an interaction between months relative to law change and and a zip code's median income ('0,000s). All regressions include zip code fixed-effects and year by month fixed-effects. All standard errors are clustered at the zip-code level. Only zip codes in Washington state with greater than 5,000 population and which matched a zip code tabulation area are included. Standard errors are clustered at the zip code level.

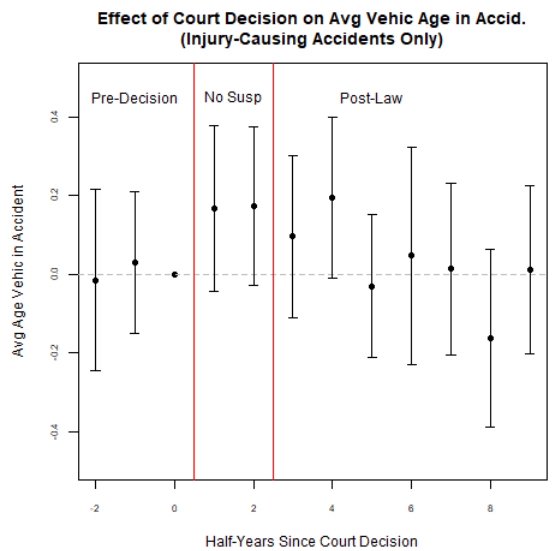


Figure 2.11: Event Study: Effect on Average Age of Vehicles in Accident by Policy Regime

Figure shows effect of the license suspension policy on the average age of vehicles involved in an accident in Washington State and Northern California. All coefficients come from regressions with jurisdiction by county fixed effects and agency type by date fixed effects. All standard errors are clustered at the jurisdiction level. Sample includes all accidents between January 2003 and June 2008. Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS). Standard errors are clustered at the jurisdiction level.

Table 2.1: Placebo Test: Prior Offenses in 18 Months Before Offense

VARIABLES	(1) Any DWLS	(2) Ttl DWLS	(3) Any Oth Traf	(4) Any Oth Traf	(5) Any Non-Traf Misd	(6) Ttl Non-Traf Misd	(7) All Cases Dism
New Law	-0.000139	-0.0104	0.00218	0.00150	0.00139	-0.00638	0.00823
	(0.00715)	(0.0144)	(0.00737)	(0.0147)	(0.00739)	(0.0123)	(0.00676)
Sample	All	All	All	All	All	All	All
Dep. Var. Mean	0.0988	0.151	0.0859	0.139	0.0949	0.132	0.124
Obs Left	49139	49128	49139	49101	49139	49080	49139
Obs Right	49349	49321	49349	49316	49349	49295	49349

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on various measures of criminal history in the year and a half prior to the violation. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts. Values greater than the 99th percentile of non-zero outcomes were excluded as outliers for all non-binary outcomes, as the distributions are heavily right-skewed.

Table 2.2: Placebo Test: Punishment for Original Offense and Demographic Characteristics

VARIABLES	(1) Total Unsusp. Fine	(2) Sentenced Fine	(3) Ttl Jail Time Srvd	(4) Total Jail Time Sent.	(5) Ln Zip Med Inc	(6) Ln LEA Med Inc
New Law	8.168	-48.98	-0.552	-0.849	-0.00405	0.0100
	(15.28)	(49.38)	(1.368)	(3.712)	(0.00601)	(0.00939)
Sample	All	All	All	All	All	All
Dep. Var. Mean	758.1	2006	23.22	216.1	43799	42801
Obs Left	48797	49042	48916	49091	46539	19855
Obs Right	48950	49139	49090	49266	46686	21188

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on sentenced fine, total fine, sentenced jail days, served jail days, zip code median income and law enforcement agency service territory median income. For fines and jail days, values greater than the 99th percentile were excluded as outliers. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts. Values greater than the 99th percentile of non-zero outcomes were excluded as outliers for all non-binary outcomes, as the distributions are heavily right-skewed.

Table 2.3: Effects of License Suspension Threat on Outstanding FTAs

VARIABLES	(1) Any FTA (2Yrs)	(2) Any FTA (4Yrs)	(3) Any FTA (6Yrs)	(4) Any FTA (8Yrs)	(5) Any FTA (10Yrs)
New Law	-0.0302	-0.0397	-0.0499	-0.0532	-0.0519
	(0.00926)	(0.00925)	(0.00915)	(0.00889)	(0.00867)
Sample	All	All	All	All	All
Dep. Var. Mean	0.268	0.268	0.258	0.242	0.227
Obs Left	49139	49139	49139	49139	49139
Obs Right	49349	49349	49349	49349	49349

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on the existence of an outstanding FTA various time periods after the original violation. A violation is considered to have an outstanding FTA if an FTA is open at the date shown (i.e if two years, then is considered to have an FTA if FTA is open exactly two years after violation). FTAs are issued for failure to appear in court, failure to respond to a citation or failure to comply with a punishment. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts.

Table 2.4: Effects of License Suspension Threat on Fine Repayment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Paid Amt	Paid Amt	Paid Amt	Paid Amt	Any Paid	Any Paid	Any Paid	Any Paid
New Law	67.15	69.18	77.57	-4.338	0.0330	0.0332	0.0737	0.0109
	(13.58)	(11.93)	(15.99)	(4.050)	(0.00897)	(0.00826)	(0.0179)	(0.00683)
Sample	All	All	Outliers	Non-Traf Misd	All	All	All	Non-Traf Misd
Bandwidth	1 yr	Optimal	1 yr	1 yr	1 yr	Optimal	3 mths	1 yr
P	2	2	2	2	2	2	2	2
Kernel	Tri	Tri	Tri	Tri	Tri	Tri	Tri	Tri
Dep. Var. Mean	666.5	666.5	666.5	163.5	0.755	0.755	0.755	0.464
Obs Left	48429	64281	49139	95650	49139	57750	12314	95714
Obs Right	48445	64378	49349	98712	49349	58184	12263	98822

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on the amount and probability of payment by May 2018. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses except where indicated. All case data come from the Washington Administrative Office of the Courts. Except where indicated, values greater than the 99th percentile of non-zero outcomes were excluded as outliers for all non-binary outcomes, as the distributions are heavily right-skewed.

Table 2.5: Effects of License Suspension Threat on FTA-Related License Suspensions

VARIABLES	(1) Any Susp (2Yrs)	(2) Any Susp (2Yrs)	(3) Any Susp (2Yrs)	(4) Any Susp (2Yrs)	(5) Any Susp (2Yrs)
New Law	0.242	0.243	0.241	0.245	0.241
	(0.00683)	(0.00467)	(0.00566)	(0.0134)	(0.00641)
Sample	Outliers	All	All	All	All
Bandwidth	12 mths	12 mth	Optimal	2 mths	12 mths
P	2	1	2	2	2
Kernel	Tri	Tri	Tri	Tri	Uniform
Dep. Var. Mean	0.00117	0.00117	0.00117	0.00117	0.00117
Obs Left	43862	43862	63985	10703	43862
Obs Right	43967	43967	64457	10659	43967

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on the probability that a license suspension for FTA (failure to appear, failure to comply or failure to respond) is attached to a case ten years after the original violation. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts and all license data come from the Washington Department of Licensing (DoL). License suspensions and case data were merged using citation numbers; approximately 22% of cases have duplicate citations numbers and so were dropped from the sample. Only suspensions that began two years or less following the offense were included.

Table 2.6: Effects of Court Decision on Average Age of Vehicles in Accidents: Any Evident Injuries

VARIABLES	(1) Vehic Age	(2) Vehic Age	(3) Vehic Age	(4) Vehic Age	(5) Vehic Age	(6) Vehic Age
WA x Court Decision Active	0.157		0.143		0.125	
	(0.0564)		(0.0692)		(0.0953)	
WA x Pre-Decision		-0.187		-0.141		-0.163
		(0.0592)		(0.0856)		(0.107)
WA x Post Law Yr 1		-0.0326		0.0406		-0.0338
		(0.0648)		(0.0996)		(0.108)
WA x Post Law Yr 2		-0.174		-0.229		-0.121
		(0.0817)		(0.0918)		(0.145)
WA x Post Law Yr 3		-0.243		-0.292		-0.179
		(0.0823)		(0.0981)		(0.119)
Observations	220,691	220,691	208,284	208,284	113,277	113,277
R^2	0.058	0.058	0.183	0.183	0.233	0.234
Sample	All	All	All	All	Muni-only	Muni-only
Spec	Base	Base	Cnty Char x Date FE	Cnty Char x Date FE	Muni Char x Date FE	Muni Char x Date FE
Dep. Var. Mean	8.239	8.239	8.239	8.239	8.228	8.228

Robust standard errors in parentheses

Table shows the effect of the license suspension policy on the age of vehicles involved in an accident in Washington State and Northern California. All regressions include jurisdiction by county fixed effects and agency type by date fixed effects. Other controls and sample restrictions are specified. Only accidents with evident injuries were included. Baseline-characteristics include county poverty decile and poverty population decile for the full sample and municipal poverty decile and municipal population decile for the municipal sample. All standard errors are clustered at the jurisdiction level. Sample includes all injury-producing accidents between January 2003 and June 2008. “Court Decision Active” refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

Table 2.7: Effects of Court Decision on Crash Rates in Accidents with Any Evident Injuries by Jurisdiction Poverty Status

VARIABLES	(1) Crashes/1KPop	(2) Crashes/1KPop	(3) Crashes/1KPop	(4) Crashes/1KPop	(5) Crashes/1KPop	(6) Crashes/1KPop	(7) Crashes/1KPop	(8) Crashes/1KPop
Court Decis Active x WA x > Med Pov	0.166						0.143	
	(0.0771)						(0.0695)	
Pre-Decision x WA x >Med Pov		-0.197						-0.128
		(0.106)						(0.0853)
Post Law Yr 1 x WA x >Med Pov		-0.144						-0.0681
		(0.0789)						(0.0758)
Post Law Yr 2 x WA x >Med Pov		-0.215						-0.244
		(0.0976)						(0.0925)
Post Law Yr 3 x WA x >Med Pov		-0.0955						-0.139
		(0.131)						(0.118)
Court Decis Active x WA x Pov Share			0.647					
			(0.497)					
Pre-Decision x WA x Pov Share				-0.846				
				(0.612)				
Post Law Yr 1 x WA x Pov Share				-0.640				
				(0.437)				
Post Law Yr 2 x WA x Pov Share				-0.879				
				(0.642)				
Post Law Yr 3 x WA x Pov Share				-0.141				
				(0.860)				
Court Decis Active x > Med Pov					0.189			
					(0.0667)			
Pre -Decision x > Med Pov						-0.152		
						(0.0899)		
Post Law Yr 1 x > Med Pov						-0.161		
						(0.0688)		
Post Law Yr 2 x > Med Pov						-0.242		
						(0.0804)		
Post Law Yr 3 x > Med Pov						-0.217		
						(0.112)		
Observations	14,048	14,048	14,048	14,048	6,062	6,062	19,591	19,591
R^2	0.806	0.806	0.807	0.807	0.770	0.771	0.706	0.706
Spec	Triple-Dif	Triple-Dif	Triple-Dif	Triple-Dif	Dif-in-Dif	Dif-in-Dif	Triple-Dif	Triple-Dif
Sample	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only
Pop Cut-off	5k	5k	5k	5k	5k	5k	1k	1k
Dep. Var. Mean	1.252	1.252	1.252	1.252	1.518	1.518	1.252	1.252

Clustered standard errors in parentheses

Table shows effect of the license suspension policy on the crash rate of vehicles in injury-producing accidents in Washington State and Northern California by poverty status. Above median poverty rate refers to all jurisdictions above the median poverty rate for our sample (10.3 percent). All regressions include jurisdiction fixed effects, state by year-month fixed-effects and poverty indicator (or continuous variable) by year-month fixed-effects. Other controls and sample restrictions are specified. All standard errors are clustered at the jurisdiction level. Sample includes all injury-producing accidents between January 2003 and June 2008. "Court Decision Active" refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

B.1 Tables and Figures

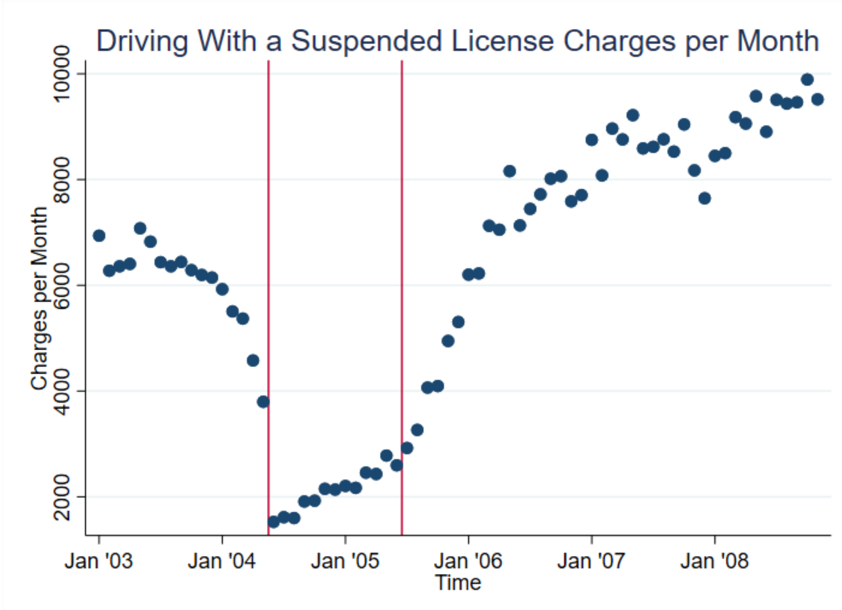


Figure B.1: Driving With a Suspended Licenses by Month

Figure shows total Driving with License Suspended (DWLS) cases in the state of Washington by month between 2003 and 2008. Data come from the Washington Administrative Office of the Courts.

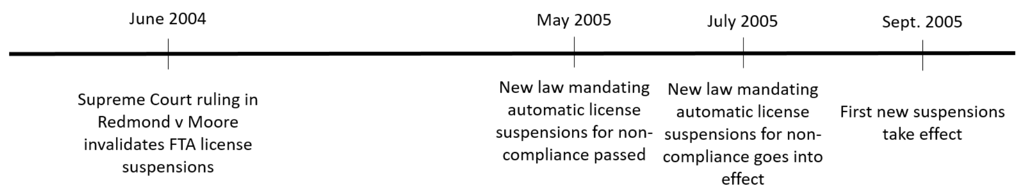


Figure B.2: Policy Timeline

This figure shows the timeline of events relating to license suspension laws in Washington state over the time period studied.

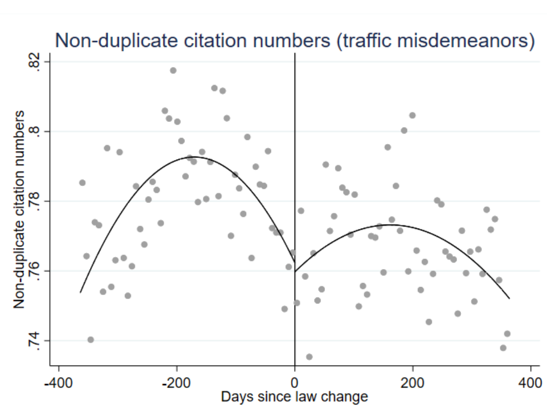


Figure B.3: Examination of Bias in Citation Matching Procedure

Figure shows the effect of the discontinuity on whether or not a case has a duplicate citation number. Outcomes are binned at the weekly level. The sample consists of traffic misdemeanor cases with at least one non-license misdemeanor charge. All case data come from the Washington Administrative Office of the Courts and all license data come from the Washington Department of Licensing (DoL).

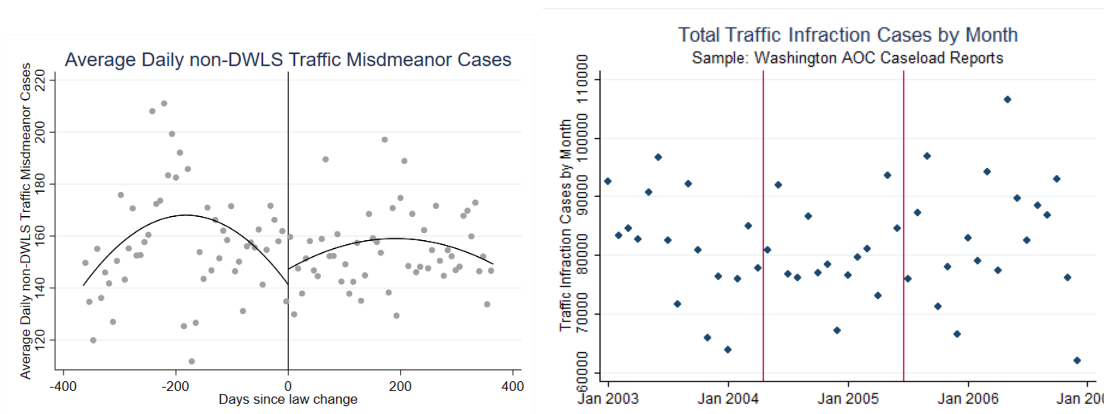


Figure B.4: Number of Traffic Misdemeanor Cases (biweekly) and Traffic Infractions (monthly) Relative to Discontinuity

Figure shows total number of traffic misdemeanor (left) and traffic infraction cases (right). Traffic misdemeanor cases are from the charge-level records provided by the Washington AOC and are aggregated at the weekly-level. Only cases with at least one non-license offense are included. Traffic infraction cases are from the AOC monthly caseload archive and are provided at the monthly level.

Figure B.5: Changes in Charges, Fines and Sentenced Days by Court by Case Type

The figure shows number of charges, number of sentenced days and total fines by court by quarter associated with license misdemeanors and non-traffic misdemeanors. The grey dashed line represents the law change reinstating license suspensions.

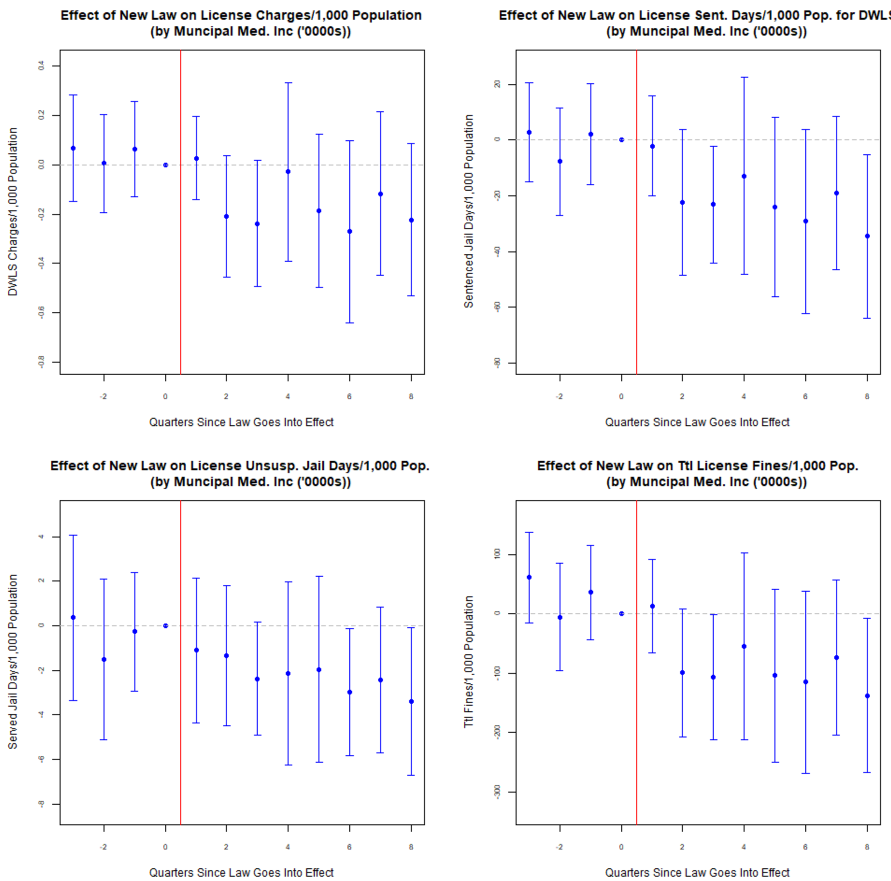


Figure B.6: Effect of New Law on DWLS by Law Enforcement Agency Service Area Median Income

The figure shows the coefficients from a regression of zip code monthly license charges per 1,000 population, license misdemeanor fines per 1,000 population, license misdemeanor sentenced jail days per 1,000 population and license misdemeanor served days per 1,000 population on an interaction between months relative to law change and a municipalities' median income ('0,000s). All regressions include law-enforcement agency fixed-effects and year by month fixed-effects. Only law enforcement agencies in Washington state with greater than 5,000 population are included. All standard errors are clustered at the law enforcement agency level.

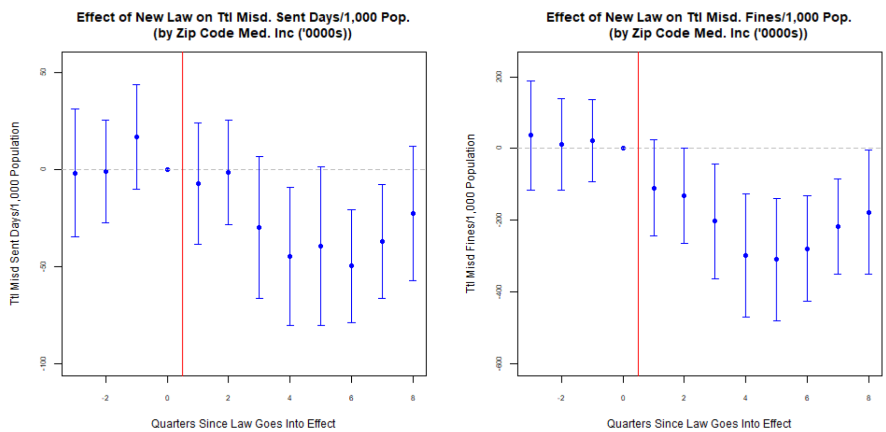


Figure B.7: Effect of New Law on Total Misdemeanor Sentenced Jail Days and Fines by Zip Code Median Income

The figure shows the coefficients from a regression of zip code monthly misdemeanor fines per 1,000 population and misdemeanor sentenced jail days per 1,000 population on an interaction between months relative to law change and and a zip code's median income ('0,000s). All regressions include zip code fixed-effects and year by month fixed-effects. All standard errors are clustered at the zip-code level. Only zip codes in Washington state with greater than 5,000 population and which matched a zip code tabulation are are included.

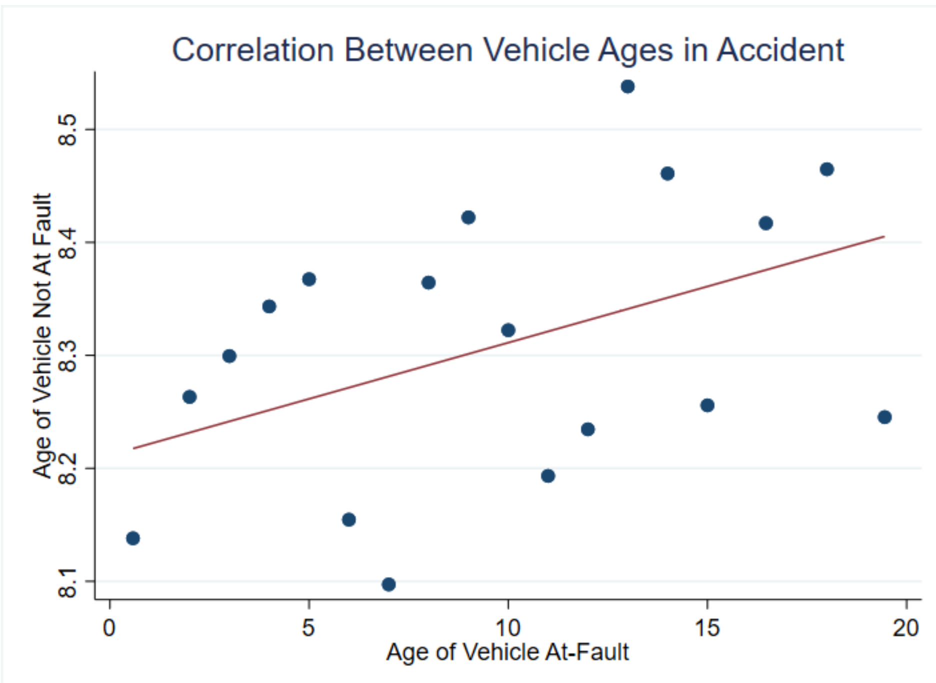


Figure B.8: Correlation Between Ages of Vehicles in Accident

Figure shows the correlation between the age of the vehicle at fault in the accident and the age of the vehicle not at fault in an accident. Only accidents in Washington state between 2003 and 2008 with evident injuries were included. Data are from the Washington Department of Transportation.

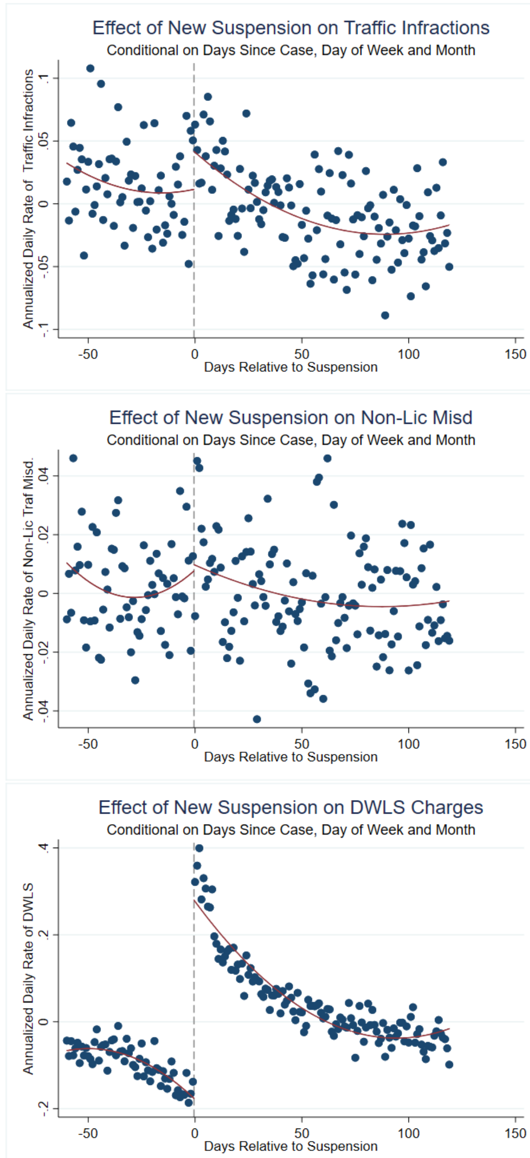


Figure B.9: Changed in Annualized Probability of Various Offenses Around Suspension Date

Figure shows effect of a license suspension on the annualized probability (daily rate of offense*365) of moving traffic infractions (top), non-license traffic misdemeanors (middle) or driving with license suspended offenses (bottom). Only license suspensions for FTAs that occurred between 2014 and 2017 are used. Less than 2% of cases with duplicate citation numbers were dropped. All outcome variables are residuals from a regression of the outcome on month, day-of-week and days since a case was filed. Data on suspensions comes from the Washington Department of Licensing and data on case filing and outcomes comes from the Washington Administrative Office of the Courts.

Table B.1: Placebo Check: Effect of 2006 or 2007 "Law Change"

VARIABLES	(1) FTA 10 Yr	(2) FTA 10 Yr	(3) Paid Amt	(4) Paid Amt	(5) Susp 10 Yr	(6) Susp 10 Yr
New Law	0.00593	0.00369	-7.934	-0.145	0.00850	-0.0138
	(0.00854)	(0.00819)	(13.62)	(12.70)	(0.00962)	(0.00910)
Placebo Date	7/1/2006	7/1/2007	7/1/2006	7/1/2007	7/1/2006	7/1/2006
Dep. Var. Mean	0.220	0.217	344.9	348.3	0.253	0.249

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on the probability of a FTA 10 years after the violation date, total amount repaid by 2018 and the probability of a license suspension for FTA is attached to a case at any time in the 2 years after the original violation using two placebo law dates: July 1st 2006 and July 1st 2007. All coefficients come from a local quadratic regression with triangular weights and a one-year bandwidth. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts and all license data come from the Washington Department of Licensing (DoL). License suspensions and case data were merged using citation numbers; approximately 22% of cases have duplicate citations numbers and so were dropped from the sample.

Table B.2: Effects of License Suspension Threat on Outstanding FTAs: Robustness Check

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Any FTA (10Yrs)	Any FTA (10Yrs)	Any FTA (10Yrs)	Any FTA (10Yrs)	Any FTA (10Yrs)
New Law	-0.0512	-0.0500	-0.0785	-0.0557	-0.0511
	(0.00591)	(0.00887)	(0.0177)	(0.00809)	(0.00866)
Sample	All	All	All	All	All
Bandwidth	1 Yr	Optimal	3 mths	12 mths	12 mths
P	1	2	2	2	2
Kernel	Tri	Tri	Tri	Uniform	Tri
Dep. Var. Mean	0.227	0.227	0.227	0.227	0.227
Obs Left	49139	47226	12044	49139	49139
Obs Right	49349	47320	12018	49349	49349

Clustered standard errors in parentheses

Table shows the effect of the discontinuity on the existence of an outstanding FTA ten years after the original violation. A violation is considered to have an outstanding FTA if an FTA is open at exactly ten years following the violation. Specifications, weights and bandwidths are as indicated. Standard errors are clustered at the defendant level. The sample consists of traffic misdemeanor cases with at least one charge unrelated to suspended or invalid licenses. All case data come from the Washington Administrative Office of the Courts.

Table B.3: Effects of Court Decision on Average Vehicle Value in Accidents

VARIABLES	(1) Vehic Val	(2) Vehic Val	(3) Vehic Val	(4) Vehic Val	(5) Vehic Val	(6) Vehic Val
WA x Court Decis Active	-151.4		-160.6		-340.7	
	(72.63)		(118.8)		(148.7)	
WA x Pre-Decision		185.3		180.6		354.4
		(84.82)		(148.0)		(173.1)
WA x Post Law Yr 1		4.977		-69.62		252.4
		(90.64)		(138.6)		(166.9)
WA x Post Law Yr 2		169.0		208.2		192.0
		(93.56)		(139.3)		(224.4)
WA x Post Law Yr 3		255.1		375.4		590.2
		(105.6)		(151.4)		(172.8)
Observations	200,259	200,259	188,073	188,073	101,317	101,317
R^2	0.057	0.057	0.184	0.184	0.241	0.242
Sample	All	All	All	All	Muni-only	Muni-only
Spec	Base	Base	Base	Base	Bl Char x Date FE	Bl Char x Date FE
Dep. Var. Mean	14262	14262	14262	14262	14230	14230

Clustered standard errors in parentheses

Table shows effect of the license suspension policy on the value of vehicles involved in an accident in Washington State and Northern California. All regressions include jurisdiction by county fixed effects and agency type by date fixed effects. Other controls and sample restrictions are specified. Only accidents with evident injuries were included. Baseline-characteristics include county poverty decile and poverty population decile for the full sample and municipal poverty decile and municipal population decile for the municipal sample. All standard errors are clustered at the jurisdiction level. Sample includes all accidents between January 2003 and June 2008. "Court Decision Active" refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

Table B.4: Effects of Court Decision on Average Vehicle Age in Accidents: By Accident Severity

VARIABLES	(1) Vehic Age	(2) Vehic Age	(3) Vehic Age	(4) Vehic Age	(5) Vehic Age	(6) Vehic Age	(7) Vehic Age	(8) Vehic Age
WA x Court Decis Active	0.525		0.332		0.0609		0.0959	
	(0.397)		(0.137)		(0.0302)		(0.0440)	
WA x Pre-Decision		-1.023		-0.376		-0.0459		-0.136
		(0.473)		(0.179)		(0.0347)		(0.0454)
WA x Post Law Yr 1		-0.258		-0.203		0.0279		-0.0529
		(0.486)		(0.180)		(0.0381)		(0.0645)
WA x Post Law Yr 2		-0.793		-0.479		-0.0586		0.0137
		(0.561)		(0.183)		(0.0429)		(0.0658)
WA x Post Law Yr 3		0.617		-0.243		-0.192		-0.195
		(0.743)		(0.172)		(0.0410)		(0.0596)
Observations	6,498	6,498	36,806	36,806	754,567	754,567	443,407	443,407
R^2	0.372	0.373	0.168	0.168	0.035	0.035	0.093	0.093
Sample	All	All	All	All	All	All	Muni	Muni
Spec	Base	Base	Base	Base	Base	Base	Full	Full
Inj Type	Fatal	Fatal	Serious+	Serious+	Possible+	Possible+	Possible+	Possible+

Clustered standard errors in parentheses

Table shows effect of the license suspension policy on the age of vehicle in an accident in Washington State and Northern California. Accident injury severity cut-offs are as indicated. All regressions include jurisdiction by county fixed effects and agency type by date fixed effects. Other controls and sample restrictions are specified. Baseline-characteristics include county poverty decile and poverty population decile for the full sample and municipal poverty decile and municipal population decile for the municipal sample. All standard errors are clustered at the jurisdiction level. Sample includes all accidents between January 2003 and June 2008. "Court Decision Active" refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

Table B.5: Effects of Court Decision on Average Age of Vehicles in Accidents: All California Control

VARIABLES	(1) Vehic Age	(2) Vehic Age	(3) Vehic Age	(4) Vehic Age	(5) Vehic Age	(6) Vehic Age
WA x Court Decis. Active	0.0512		0.143		0.125	
	(0.0464)		(0.0692)		(0.0953)	
WA x Pre-Decision		-0.208		-0.141		-0.163
		(0.0480)		(0.0856)		(0.107)
WA x Post Law Yr 1		0.0916		0.0406		-0.0337
		(0.0529)		(0.0996)		(0.108)
WA x Post Law Yr 2		0.0265		-0.229		-0.121
		(0.0678)		(0.0918)		(0.145)
WA x Post Law Yr 3		-0.0540		-0.292		-0.179
		(0.0681)		(0.0981)		(0.119)
Observations	583,028	583,028	208,284	208,284	113,282	113,282
R^2	0.042	0.042	0.183	0.183	0.233	0.234
Sample	All	All	All	All	Muni-only	Muni-only
Spec	Base	Base	Cnty Char x Date FE	Cnty Char x Date FE	Bl Char x Date FE	Bl Char x Date FE

Clustered standard errors in parentheses

Table shows effect of the license suspension policy on the age of vehicles involved in an accident in Washington State and California. All regressions include jurisdiction by county fixed effects and agency type by date fixed effects. Other controls and sample restrictions are specified. Only accidents with evident injuries were included. Baseline-characteristics include county poverty decile and poverty population decile for the full sample and municipal poverty decile and municipal population decile for the municipal sample. All standard errors are clustered at the jurisdiction level. Sample includes all accidents between January 2003 and June 2008."Court Decision Active" refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

Table B.6: Effects of Court Decision on Injury Rates in Accidents with Any Evident Injuries by Jurisdiction Poverty Status

VARIABLES	(1) Crashes/1KPop	(2) Crashes/1KPop	(3) Crashes/1KPop	(4) Crashes/1KPop	(5) Crashes/1KPop	(6) Crashes/1KPop	(7) Crashes/1KPop	(8) Crashes/1KPop
Court Decis. Active x WA x > Med Pov	0.0290						0.0250	
	(0.0259)						(0.0230)	
Pre Decis x WA x > Med Pov		-0.0235						-0.0204
		(0.0378)						(0.0306)
Post Law Yr 1 x WA x > Med Pov		-0.0288						-0.0187
		(0.0312)						(0.0342)
Post Law Yr 2 x WA x > Med Pov		-0.0364						-0.0494
		(0.0335)						(0.0289)
Post Law Yr 3 x WA x > Med Pov		-0.0296						-0.0135
		(0.0359)						(0.0323)
Court Decis. Active x WA x Pov Share			0.0687					
			(0.155)					
Pre-decision x WA x Pov Share				-0.0755				
				(0.203)				
Post Law Yr 1 x WA x Pov Share				-0.0919				
				(0.178)				
Post Law Yr 2 x WA x Pov Share				-0.0573				
				(0.203)				
Post Law Yr 3 x WA x Pov Share				-0.0472				
				(0.235)				
Court Decis. Active x WA					0.0590			
					(0.0226)			
Pre-Decision x WA						-0.0311		
						(0.0325)		
Post Law Yr 1 x WA						-0.0607		
						(0.0271)		
Post Law Yr 2 x WA						-0.0724		
						(0.0283)		
Post Law Yr 3 x WA						-0.0835		
						(0.0313)		
Observations	14,058	14,058	14,058	14,058	6,072	6,072	19,602	19,602
R^2	0.598	0.598	0.598	0.598	0.574	0.574	0.447	0.447
Spec	Triple-Dif	Triple-Dif	Triple-Dif	Triple-Dif	Dif-in-Dif	Dif-in-Dif	Triple-Dif	Triple-Dif
Sample	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only	Muni-only
Pop Cut-off	5k	5k	5k	5k	5k	5k	1k	1k
Dep. Var. Mean	0.307	0.307	0.307	0.307	0.342	0.342	0.307	0.307

Robust standard errors in parentheses

Table shows effect of the license suspension policy on the injury rate of vehicles in injury-producing accidents in Washington State and Northern California by poverty status. Above median poverty rate are all jurisdictions with greater than a 10.3 percent poverty rate. All regressions include jurisdiction fixed effects, state by year-month fixed-effects and poverty indicator (or continuous variable) by year-month fixed-effects. Other controls and sample restrictions are specified. All standard errors are clustered at the jurisdiction level. Sample includes all injury-producing accidents between January 2003 and June 2008. “Court Decision Active” refers to the period between June 2004 and July 2005 and is the omitted category in Columns (2),(4), and (6). Crash data are from the Washington Department of Transportation and the California Statewide Integrated Traffic Reporting System (SWITRS).

B.2 Vehicle Valuation Estimation

To estimate the effect of the license suspension policy on traffic safety, it was necessary to impute values for vehicles involved in traffic accidents. Unfortunately, historical data on used car vehicles are not readily available, so I instead created depreciation curves based on contemporaneous (as of 2017) new and used car values by vehicle make. Specifically, I used data on vehicle valuations from a full scrape of the truecars.com website on September 24, 2017, a large online advertiser of used-car sales. Data were taken from a dataset publicly posted on kaggle.com, a data science website and various robustness tests were performed to test whether the available data matched publicly available estimates of relative brand values and depreciation scales. More than 1.2 million sales listings were included.

The Washington crash data includes information on the car make, model, year while the California data includes data on make and year. To standardize across states, I created average prices for each vehicle make x age cell for all ages up to 20 years old. I dropped the following discontinued makes: Saturn, Plymouth, Mercury, Oldsmobile and Saab as well as any cell years with fewer than twenty sales. For the remaining makes, I interpolated (or extrapolated) any missing age cells using log price to reflect the log-linear relationship between price and age (and then exponentiating so as to return the outcome variable to levels). I then used these variables to create depreciation curves for each make. Using this balanced panel of make x vehicle age cells, I then merged with crash-level data for Washington and California. The underlying assumption is that the relative values and depreciation curves of each make is similar in the 2017 period and the 2004-2008 study period. While this procedure certainly creates measurement error, so long as this error is uncorrelated with the court decision and subsequent law change, it will only increase the standard errors of our estimates for my primary traffic accident analysis (comparing Washington and Northern California) and not introduce bias because I am using the price as a left-hand side variable.

Chapter 3

Litigation as a Policy Instrument: The Case of the New Source Review Litigation

3.1 Abstract

Enforcing regulations through litigation against noncompliant firms is an important policy tool for environmental regulators, but there is limited empirical evidence on the effectiveness of this mechanism. In this paper, I study this question by examining the effects of a major environmental enforcement initiative, which led one-third of the US coal-fired power plant fleet to come under a consent decree. I show that legal settlements arising out of this initiative did indeed lead to large decreases in plant pollution emissions. These decreased emissions further led to meaningful improvements in local air quality and decreases in local cardiovascular and respiratory mortality rates. I then conclude by showing suggestive evidence that in regulated electricity markets average electricity retail price and utility revenues increased following a settlement.

3.2 Introduction

In many areas of administrative law, the definition of a violation depends upon interpretations of often ambiguously-worded rules. As such, even in the absence of changes in existing rules, regulations can be made more or less strict through strategic decisions about enforcing borderline rule violations, often through litigation. This type of litigation, either by the government or by third parties has been an important driver of environmental regulation.[1] Yet, despite the importance of litigation-based enforcement, we know relatively little about its overall effectiveness in the environmental context.

In particular, there are two major reasons why we might expect that litigation, even if successful, may deliver only limited environmental benefits. One, both regulators and firms may have incentives to agree to changes that the firm would have undertaken even in the absence of litigation. Two, any agreed-upon changes must be enforced by the government. Because policy

[1] According to the EPA's Enforcement and Compliance History Online (ECHO) database, EPA enforcement cases originating between 2009 and 2015 have led firms to agree to compliance actions valued between $1 and $10 billion each year between 2009 and 2015.

priorities and political regimes change over time and because incentives for ensuring compliance may differ from incentives for achieving a settlement, it is not a foregone conclusion that any changes dictated in a legal settlement will occur in practice. Finally, even if litigation is successful in forcing change, the ultimate incidence of the costs of these changes are unclear.[2]

In this paper, I attempt to investigate these questions empirically by examining the effects of a major set of environmental litigation: the EPA/DOJ New Source Review (NSR) Initiative. This initiative led more than one-third of the US coal-fired power plant fleet to come under consent decrees between 2000 and 2015. These consent decrees required affected generating units to meet certain emissions standards and/or install costly pollution control equipment—actions estimated by the EPA to cost more than $20 billion.[3]

I begin by examining whether or not these settlements were successful in accomplishing their stated pollution reduction goals. Using a series of difference-in-differences designs, I first show that settlements arising out of this litigation did indeed lead to large reductions in emissions of NO_x and SO_2. Emissions of both pollutants fell by 10-20% on average and this decline was driven by changes in emission rates; there were no changes in overall levels of generation.

I next find suggestive evidence that these emission declines led to meaningful changes in local ambient air quality and mortality. Specifically, using air pollution monitor data I observe that ambient sulfur dioxide levels fell by more than 10% around affected plants after the settlement was in place with suggestive evidence for additional smaller declines in PM 2.5. I then show evidence that cardiovascular and respiratory mortality fell significantly in counties near plants under NSR consent decrees; having a plant come under consent decree within 40 miles of a county's centroid lead to a 3.5 deaths per 100,000 population decline in cardiovascular-related deaths and 2.2 deaths per 100,000 population decline in respiratory-related deaths. There were no

[2]For instance, regulated utilities may be expected to recoup any capital expenses forced by policy, but the same may not be true if changes occur as a result of litigation.

[3]This number is based on the author's calculation using data from the EPA's ECHO database. It is important to note that EPA estimates are likely based on engineering estimates and so may either overestimate or underestimate the true costs to utilities.

changes in externally-caused deaths, nor changes in mortality in counties where a consent decree occurred 40-80 miles from the county centroid.

I finally analyze who bore the cost of these improvements in pollution. I show suggestive evidence that regulated investor-owned utilities (IOUs), which are typically guaranteed a minimum rate of return on their investment, increased both prices and revenues increased by more than 10% after a settlement relative to similar non-settling utilities. These results imply that at least in regulated electricity markets, utilities were able to pass on a substantial proportion of the overall compliance costs to ratepayers.

This paper contributes to several major literatures. First, it supports a growing literature on the use of the courts to enact policy in areas including education (Jackson, Johnson, and Persico, 2015), equal employment (Miller and Segal, 2012) and civil rights (Cascio and Washington, 2013). In the environmental context, there has been a growing body of work examining the specific and deterrence effects of various enforcement actions (Gray and Shimshack, 2011), as well as more recent research demonstrating the importance of non-environmental regulations for firm pollution decisions (Boomhower, 2019) and the importance of the dynamic structure of enforcement actions (Blundell, Gowrisankaran, and Langer, Forthcoming). However, the vast majority of research examining environmental enforcement has focused on either the effects of fines and other regulator penalties for rule violations or the threat of litigation; there has been relatively little evidence on the effects of major environmental litigation programs. One major contribution of this paper is to provide novel evidence showing that, at least in the air pollution context, these settlements can have large, independent effects on environmental and health outcomes.

This paper also builds on a small economics literature examining the effect of the NSR litigation itself. Keohane, Mansur, and Voynov (2009) examined whether firms responded to the threat of the onset of litigation studied here and found some evidence that targeted firms were more likely to reduce emissions prior to the onset of litigation, while Chan and Zhou (2019) show that plants at greater risk of being sued for NSR violations experienced greater reductions in carbon-dixoide emissions as well as SO_2 and NO_x. Other work by legal scholars have provided an overview of the regulatory issues at stake in this litigation, but have not examined the effects of the litigation in a causal framework (Nash and Revesz, 2007; McGarity, 2012). This paper builds on this body of work by providing novel causal evidence of the pollution-reducing effects of the NSR litigation settlements themselves.

This paper additionally adds to a new literature examining the causes of the extraordinary decline in pollution over the past two decades in the United States (Shapiro and Walker, 2018; Holland et al., 2018) as well as the effect of that decline on local air quality and health outcomes (Barreca, Neidell, and Sanders, 2017; Johnson, LaRiviere, and Wolff, 2017). In particular, it complements other estimates (Bushnell and Wolfram, 2012; Heutel, 2011; Lange and Linn, 2008; List et al., 2003) of the impact of grandfathering old plants into the Clean Air Act regulations on local pollution. I show that NSR settlements alone were an important reason that power plant SO_2 and NO_x pollution rates declined over the past two decades. These results further imply that the existence of NSR grandfathering had a large and persistent effect on overall pollution levels.

Finally, this paper contributes to a growing literature on the interaction between restructured electricity markets and environmental protection. Both Fowlie (2010) and Cicala (2015) found that plants in states with restructured electricity markets were less likely to invest in more capital-intensive pollution control technology. In this paper, I build on this work by showing that regulated utilities were able to pass a substantial proportion of settlement costs onto ratepayers providing further evidence of an incentive to invest in pollution control in regulated markets.

The remainder of this paper is organized in the following manner. Section 2 provides background on the studied litigation. Section 3 describes the data and empirical strategy. Section 4 describes the primary results and Section 5 concludes.

3.3 Policy Background

This paper studies the effects of a series of enforcement initiatives undertaken by the EPA and Department of Justice over violations of the Clean Air Act. In the Clean Air Act (CAA,1970) and Clean Air Act Amendments (CAAA,1977), Congress created emissions standards for new sources of pollution. These New Source Performance Standards (NSPS) required all major new sources *or modified existing sources* of pollution that commenced construction after a certain date meet emissions thresholds for major pollutants.[4] Additionally, after 1977 plants built or modified in NAAQS non-attainment areas were required to install control technology to meet the lowest-achievable emissions rate (LAER), while those in attainment areas were required to install the best available control technology (BACT). Both rules led to large reductions in emissions among affected units. However, a large proportion of the infrastructure used in America's most-polluting industries was built prior to these rule changes. For instance, units built prior to 1971 supplied more than half of the United States' coal-fired energy generation in 1998, while 85% of current operating US petroleum refineries were built prior to 1975. As a result, these new construction requirements did not bind for many highly-polluting industries.

Yet, because the rules also applied to major modifications at existing plants, it was expected that plants grandfathered in to the new rules would soon be forced to comply as they either undertook "major modifications" to extend the life of the plant or retired. However, the definition of what constituted a major modification was ambiguous, especially with respect to routine maintenance operations or changes that increased the plant's overall efficiency. The EPA generally held that any modifications that increased the total annual emissions of a plant were classified as major, an interpretation that was upheld in *Wisconsin Elec. Power Co. v. Reilly*

[4]Specifically, the Clean Air Act stated that these performance standards would apply to: "any stationary source, the construction or modification of which is commenced after the publication of regulations (or, if earlier, proposed regulations) prescribing a standard of performance under this section which will be applicable to such source" where modification is defined as "any physical change in, or change in the method of operation of, a stationary source which increases the amount of any air pollutant emitted by such source or which results in the emission of any air pollutant not previously emitted." (*Wisconsin Elec. Power Co. v. Reilly* 1990)

(1990) and *Environmental Defense v. Duke Energy Corporation* (2007).[5] After the ruling in *Wisconsin Elec. Power Co. v. Reilly* (1990), the EPA expected that plants would begin applying for New Source Review permits for many actions that would lead to increased plant utilization (and by extension, increased emissions) such as major maintenance and efficiency-enhancing changes. However, in the years following this case despite a boom in electricity and industrial production, relatively few permit applications were received (Nash and Revesz, 2007).

In response, the EPA began enforcing New Source Review through litigation, with separate initiatives against the petroleum refining industry, the cement industry, the glass industry and the focus of this paper, the coal-fired power plant industry.[6] In October 1999, the EPA and Department of Justice brought suit against seven utilities for violations of New Source Review. Violations typically consisted of efficiency-enhancing modifications that allowed for more intensive use of the unit and/or extended its life and that often occurred many years prior to the suit. Over the next decade, suits against more than thirty additional utilities and over a hundred plants continued to be filed. Almost all sued firms ultimately signed consent decrees ($>80\%$).

Table 3.1 shows differences across key variable for coal plants involved in litigation and settling at different times, while Table C.2 shows these same associations, but conditioning on state fixed effects so as to parallel the main empirical specifications used in the paper. Unsurprisingly, in both cases we see that coal units that ever settle litigation use more heat input at baseline (1998) than non-settlers. Plants that are sued also appear to have (marginally) significantly higher baseline emission rates, although such differences shrink dramatically when including state fixed effects. Examining timing of settlement, we see that earlier settlements are associated with higher baseline emission rates and older plants, but again after controlling for state fixed effects these associations, with the exception of 1998 NO_x emission rate, shrink dramatically. In the next section, I discuss various empirical strategies for identifying the effects of the NSR

[5]Many utilities argued both before and after these rulings that an hourly standard should instead hold—as long as the average hourly emissions of a unit did not increase, a major modification had not occurred.

[6]Data limitations prevent a detailed examination of the effects on other affected industries.

settlements and overcoming the potential biases created through these underlying differences between affected and unaffected plants.

The left panel of Figure 3.1 shows the proportion of 1998 sulfur dioxide production that was under litigation and signed onto a consent decree for the years between 1999 to 2016. Coal units responsible for more than 40% of 1998 sulfur dioxide production ultimately signed consent decrees and these settlements were staggered over time. These consent decrees typically imposed strong pollution control requirements on covered plants. McGarity (2012), in his overview of the NSR litigation, notes that these consent decrees all "contained similar features, including a denial of liability by the defendants, retirement or installation of control technologies on existing units, a prohibition on selling or trading any excess emissions allowances from controls required by the consent decrees (to ensure that the emissions reductions actually occurred), relatively low fines, a requirement to invest in supplemental environmental projects, and protection from future NSR enforcement actions for a specified number of years." Many decrees also included emission rate limits.[7]

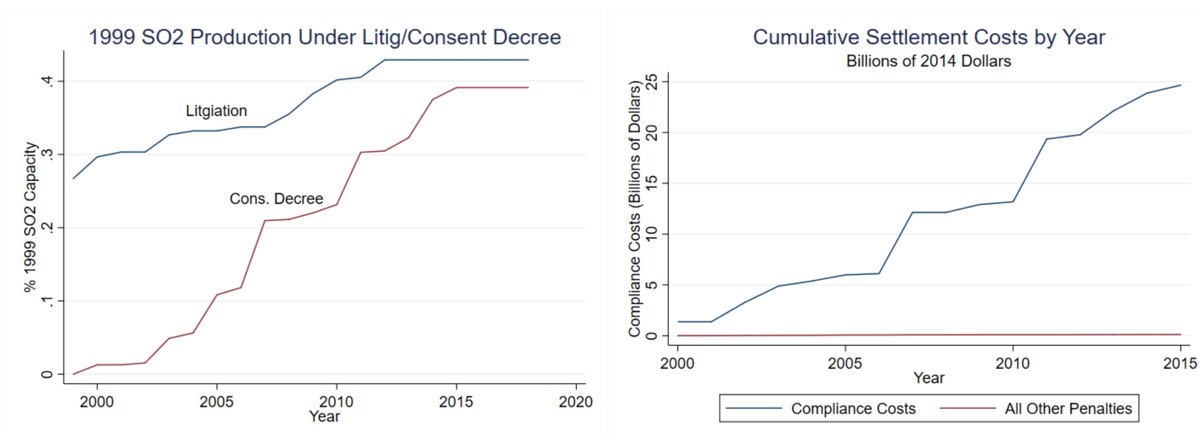
The right panel of Figure 3.1 shows the cumulative costs of these consent decrees as estimated by the EPA. Unsurprisingly, given the strict requirements embedded in the consent decrees, these settlements imposed large costs on firms; by 2016, the EPA estimated that firms had paid more than 20 billion dollars of costs through settlement agreements, but these costs almost all arose from expensive compliance actions such as the installation of pollution control equipment.[8] Other costs such as civil penalties to local, state or federal governments were negligible in comparison.

However, despite the ostensible success of the litigation, it was a priori unclear whether the requirements enumerated in the consent decrees would actually lead firms to change behavior.

[7] The prohibition against selling excess emissions permits is particularly important for understanding the general equilibrium effects of these consent decrees on overall pollution levels. Otherwise, the changes created by these consent decrees could have just led to a reallocation rather than a reduction in overall emissions.

[8] It is important to note that EPA estimates are likely based on engineering estimates and so may either overestimate or underestimate the true costs to utilities.

Many other regulations were coming into effect during this time period and firms could have agreed to compliance actions that they would have undertaken regardless of the settlement. Additionally, consent decrees must be enforced and it is not obvious that the Environmental Protection Agency and the Justice Department would have the political will or institutional capacity to do so. Finally, even if consent decrees were enforced it was unclear who would bear the burden of these settlements: ratepayers or other utility stakeholders. In the remainder of this paper, I use a variety of empirical techniques to shed light on these questions.



The left panel of this figure shows the proportion of 1998 sulfur dioxide emissions by coal-fired power plants that were under litigation or consent decree by year. The right panel of this figure shows the cumulative total compliance and noncompliance costs associated with NSR settlements by year as estimated by the EPA.

Figure 3.1: Plants Litigation Status and Total Penalties Paid by Year

Table 3.1: Association Between Treatment and Key Baseline Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Heat input	NO_x Rate	SO_2 Rate	Op Year	Heat input	NO_x Rate	SO_2 Rate	Op Year
Ever-Settled	3.497*	27.96*	116.7	-2.148				
	(1.871)	(16.44)	(91.21)	(1.503)				
Settlement Year					-0.332	-4.756*	-21.21*	-0.399*
					(0.348)	(2.455)	(10.95)	(0.220)
Observations	1,068	1,068	1,068	1,068	350	350	350	350
R^2	0.007	0.008	0.008	0.007	0.005	0.018	0.017	0.026
Dep. Var Mean (1998)	18.89	289.4	794.3	1965	18.89	289.4	794.3	1965
Units	Tril. BTU	Tons/Tril BTU	Tons/Tril BTU	Years	Tril. BTU	Tons/Tril BTU	Tons/Tril BTU	Years

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of ever settling or settlement year on various baseline outcomes. There are 1,068 units included in the regression from 409 plants. In Columns (5)-(8) only units that ever settle are included—there are 350 of such units from 109 plants. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year in order to not overweight plants with many units. Standard errors are clustered at the state level.

3.4 Data and Empirical Strategy

3.4.1 Empirical Strategy

Effect on Emissions

The goal of each empirical analysis in this paper is to find a suitable counterfactual for generating units that came under consent decrees.[9] My primary analysis compares generating

[9] I define a unit at a plant to be treated if any unit at the plant came under consent decree because of the high likelihood of within-plant spillovers if some units in a plant are under a consent decree and not others. Despite this, all analyses still use the generating unit as the unit of analysis. Different units within the same plant may be subject to different regulations and have different expected trajectories in power generation and emission rates depending upon various unit characteristics such as size and existing pollution control equipment; by using the generator as my unit of analysis I can better control for these potential confounders. If instead the analysis were run at the plant level,

units before and after coming under a consent decree relative to units that are not yet under a consent decree or which were never under litigation. To control for confounding variables that may be associated with both coming under a consent decree and changing levels of emissions, all comparisons are made within narrowly defined, mutually-exclusive cells. Units are placed in the same cell if they are in the same state, all above/below the median of 1998 heat input, above/below the median of 1998 SO_2 emission rate and all above/below the median of 1998 NO_x pollution rate. I show robustness to creating cells with more granular quantiles as well as alternative approaches. To ensure that I am not overweighting plants with more units, I weight all regressions by a unit's 1998 share of its plant's overall energy use. The identifying assumption in this model is that absent a consent decree, units within the same cell would have similar trends in pollution levels over time. My main specification is then as follows:

$$Y_{pt} = \alpha_p + \beta Post_{pt} + \gamma Z_{pt} + \varepsilon_{pt}$$

where Y_{pt} is an outcome variable in unit p in year t, α_p is a generating unit fixed-effect, and $Post_{pt}$ is an indicator variable equal to one if a year is after a unit's settlement date[10] and 0 otherwise. Units that never have a consent decree have a value of zero for all years. The vector Z_{pt} is a vector of cell by year fixed effects, which ensures that I am only comparing units that are within the same cell in a given year. I also further include controls for (fully-interacted) county ozone, PM and SO_2 nonattainment status by year fixed effects to control for a unit's National Ambient Air Quality Standards (NAAQS) regulatory regime and NO_x Budget Trading Program status by year fixed effects to control for a unit's inclusion in the NO_x Budget Trading Program, both of which may influence pollution control investment decisions. Finally, ε_{pt} is a mean-zero error term. The sample includes all years between 1998 and 2014, but is robust to including only

I would be less sure I was comparing similar plants as plants that have similar mean 1998 SO_2 emission rates for example, may actually be made up of different units with very different emission rates (i.e. one plant may have two moderately clean units, one plant may have one very dirty and one very clean unit).

[10]For years in which the settlement date occurred, the post indicator is equal to 1 if the settlement occurred before July 1st.

a narrower window around a consent decree.

I also run a similar event-study specification to look at dynamics before and after the settlement comes into place. In this specification, the post variable is replaced by a vector of indicators for years relative to settlement with the indicator for 1 year prior to settlement acting as the omitted variable. I assign units in non-settling plants a value of 0 for all years. All standard errors in all models are clustered at the state level.

There are three major challenges to identification using this specification. First, there may exist unobserved factors that are correlated with both a plant's decision to enter into a consent decree and its future pollution outcomes. To better understand this threat to identification, it is helpful to understand what factors may have driven utilities to settle. One likely factor is the timing and substance of judicial rulings in each case. Almost all firms chose to fight this litigation and so settlement depended on both the speed and outcomes of rulings by local District Court and Circuit Court of Appeal judges. Because this litigation involved complex and ambiguous legal issues, judicial rulings also varied significantly across cases at both the District and Circuit Court levels, affecting the prospective value of settlement to a firm. Additionally, changing political and legal environments also likely affected firm's willingness to settle depending upon the state of the firm's case. For instance, two utilities had tentative settlements in place in 2000, which were withdrawn after the election of George W. Bush brought a more pro-business orientation to the EPA (McGarity, 2012). Similarly, American Electric Power (AEP) chose to settle after the 2007 Supreme Court ruling that upheld the EPA's interpretation of major modifications, but this decision did not affect the settlement choices of several other utilities who had won cases in District Court and were awaiting the outcome of government appeal (Clayton, 2007).[11]

The key assumption in this paper is that to the extent that the factors that influence an individual utility to settle are correlated with future pollution outcomes, they are also common

[11] As the spokesman for Alabama Power, a company under litigation, said after the suit: "Basically, the AEP decision to settle has no impact on the Alabama Power case. We won the lower court decision and the government has appealed and is challenging that decision."

to the units that serve as the settling units' counterfactuals (i.e. the other units within their cells). The specific timing of legal rulings within the confines of the NSR case are unlikely to be correlated with future pollution trends for a settling unit relative to other units in their cell, as they are plausibly orthogonal to other firm-specific factors affecting pollution outcomes around the settlement date. Similarly, while a more utility-friendly NSR ruling may indicate that a company is in a more pro-utility District or Circuit, the existence of this pro-utility orientation would not be expected to lead to a discontinuous change in unit pollution outcomes relative to counterfactual units following the settlement date. Further, to extent this ruling produced new knowledge about the "pro-utility" alignment of a given district or circuit, this knowledge should still be common to both the treated unit and its counterfactual units in the same cell (which are in the same geographies) and so should not lead to differential changes in the treated unit.

However, other potential explanations for differential settlement timing are more worrisome from an identification perspective. Most prominently, changed exposure to state and federal regulations may have influenced settlement timing if these regulations would have induced installation of the needed pollution control equipment regardless of the settlement. To address this concern, in my primary specification I control for many time-varying "common" factors that we might expect to influence a unit's expectation of future pollution status including NAAQS nonattainment status, state-level regulation and other federal regulations administered at the state-level (through state by year fixed effects) and regulations that affect plants with different pre-litigation levels of pollution (through comparing outcomes only within narrowly-defined cells based on pollution rates) to ensure that settling units are only compared to other units equally affected by these factors. I also show through event plots that there is no evidence of pollution trends prior to the signing of the consent decree and that pollution responds immediately to settlement, adding confidence that plant-specific pollution trends are not driving settlement decisions. Additionally, because settlements typically occur at the firm, not plant level, it is unlikely that for a given unit, plant-specific factors are driving settlement decisions. As a robustness check, I restrict my analysis to only plants that settled in a group and find that effects are qualitatively similar. Finally, I perform a number of additional robustness checks testing sensitivity of estimates to other specification and sample choices—in all cases estimates remain largely unchanged.

Second, because the treatment effect in the above model is estimated off of both variation in timing in settlement among plants that settle and variation in whether or not a plant settles at all, an additional challenge is that plants that are sued may have different trends from those that are not (or those that are sued earlier or later). To understand the plausibility of this identification concern, it would be helpful to know the underlying decision rules the government used in deciding which cases to pursue. Anecdotal evidence from McGarity (2012) and GAO (2012) suggests that the EPA and DOJ were capacity-constrained in their ability to undertake litigation. In Table 3.1, I show that the EPA chose to sue larger, higher-polluting plants first—a fact also found in Keohane, Mansur, and Voynov (2009) and Chan and Zhou (2019). Such a decision rule is not a problem for my identification strategy so long as conditional on the baseline covariates included, the choice of utility to be sued is uncorrelated with future pollution outcomes.

I attempt to test this assumption empirically in several ways. First, as above, event plots show no clear trends prior to settlement. Because the median time to settlement was six years after the onset of litigation, if it were truly the case that litigated firms conditional on covariates had different pollution trajectories, these trends should appear as pre-trends in my data, but I see no evidence of such effects. Second, I perform an analysis testing whether being under litigation, but not yet settling has an effect on key pollution outcomes—I find no such effect. Again, if litigated plants were on different pollution trajectories then non-litigated plants, we would not expect these differences to first show up many years later after a settlement, but instead should manifest themselves in the intervening period. Third, as a robustness check, I perform the same main specification as above, but only including plants that ever-settled so treatment variation is driven entirely by time of settlement. Results are qualitatively similar although standard errors increase substantially with the smaller sample. Finally and importantly, plants are typically sued over violations that occurred a decade or more in the past and were not clearly recognized as violations at the time, implying that sued and non-sued plants may be less different than typical comparisons between contemporaneous violators and non-violators.

A third challenge to identification is using settlement date as the event date rather than litigation date. This choice is motivated by my desire to understand the effect of a settlement alone. The median time to settlement is six years and there is a wide variation across settling firms. If I used litigation date rather than settlement date, the years since litigation coefficients would be a weighted average of the effect of litigation and the effect of settling, leading to a potentially large underestimation of the effect of settlement. However, a disadvantage of using settlement year as my event date is that if firms reacted to litigation by changing emissions patterns, I am only estimating the effect of settlement relative to being under litigation rather than the effect of settlement relative to pre-litigation norms. This concern is particularly acute given the results of Keohane, Mansur, and Voynov (2009), who found that at the onset of this litigation (plants sued in 1999), plants responded in the very short-term (1-year) by reducing emissions.[12]

I address this concern in two ways. First, if the onset of litigation changed behavior, we would then again expect to see differential trends prior to settlement as the pre-settlement period in the event study graphs encompasses the period in which litigation occurred; there is no evidence for these effects. Second, as described above, I perform a separate analysis in which I include an indicator for years in which a plant is under litigation but has not yet settled. These effects are typically very small and highly insignificant suggesting that on average sued plants are not differentially changing their behavior in response to litigation alone. These results may differ from those of Keohane, Mansur, and Voynov (2009) because I include different controls (i.e. state by year fixed effects), because I am looking at long-run responses and/or because I am looking at the full sample of litigating plants.[13]

[12]As described above, Keohane, Mansur, and Voynov (2009) and Chan and Zhou (2019) also found that plants perceived to be at greater risk of litigation responded by changing emissions, but as long as my baseline controls are identifying similar plants, this should not bias my results.

[13]It is additionally possible that even if pre-settlement litigated plants are not acting differently than non-litigated

A final concern is the potential bias created through using a two-way fixed-effect model with time-varying treatment as identified in Goodman-Bacon (2018) and Callaway and Sant'Anna (2019). In Appendix C.2, I describe an alternate empirical specification that addresses these concerns by creating a pure control group for each treated unit and a balanced panel. I show that results remain similar when using this alternative specification.

Effect on Air Quality and Mortality

I additionally examine the effect of settlements on local pollution and mortality outcomes. I cannot use exactly the same methodology here because each plant may impact pollution and mortality over a large geographic area leading air quality monitors and counties, my unit of analysis for mortality, to be potentially treated by multiple plants. Accordingly, I instead implement a modified analysis that attempts to convert the preceding analysis to the air-quality monitor (county) setting where units may be multiply-treated depending on the number of nearby plants coming under consent decree. My sample consists of all monitors (counties) within 80 miles of a coal plant. I then create indicator variables equal to 1 if a plant has come under a consent decree within 40 miles and within 40 miles to 80 miles of each monitor (county centroid) respectively. I use a 80 mile radius in my primary analysis both because many plants have tall smokestacks that disperse pollutants many miles from their source (Barreca, Neidell, and Sanders, 2017) and because the large size of many American counties means that even counties whose population centroids are many miles from the plant may still contain areas that are relatively close-by. I show robustness to using differently sized distance bins.

As above, I group together similarly exposed monitors (counties) into narrowly-defined, mutually exclusive cells. In the primary analysis, I group together monitors (counties) that are in the same state, that are all above/below the median 1998 level of heat input used by plants within

plants, instead all firms may change behavior in response to the threat of litigation. Although I cannot test for this empirically, this is consistent with the findings of both Keohane, Mansur, and Voynov (2009) and Chan and Zhou (2019) and would suggest that I am underestimating the true total effects of litigation.

40 miles, are all above/below the median of 1998 heat input used by plants within 40-80 miles and are all above and below the median 1998 value of the outcome variable (pollution levels and mortality respectively).[14] Specifically, I estimate:

$$Y_{mt} = \alpha_m + Post40_{mt} + Post80_{mt} + Z_{mt} + \varepsilon_{mt}$$

where Y_{mt} is the outcome variable of interest for monitor (county) m in year t, α_m is a monitor (county) fixed-effect, $Post40$ is an indicator equal to 1 if a consent-decree has been implemented in a plant within 40 miles of the monitor (county), $Post80$ is an indicator equal to 1 if a consent-decree has been implemented in a plant within 40 to 80 miles of the monitor (county), and Z_{mt} is a vector of indicators for cell by year fixed effects. I further include controls for (fully-interacted) county ozone, PM and SO_2 nonattainment status by year fixed effects to control for nearby units' National Ambient Air Quality Standards (NAAQS) regulatory regime and NO_x Budget Trading Program status by year fixed effects to control for nearby units' inclusion in the NO_x Budget Trading Program.[15] Finally, ε_{mt} is a mean zero error term.

This specification implicitly compares monitors (counties) near plants that experienced a consent decree before and after the consent decree occurred relative to counties near plants that did not experience (or had not yet experienced) a consent decree. The identifying assumption here is that absent the consent decree these monitors (counties) outcomes should evolve on similar paths. I test this assumption in two ways. First, as above, I create event study plots to see how monitors' (counties') outcomes evolve before and after a consent decree occurs. Second, if the effects are truly caused by the consent decree, we should expect effects to be larger for monitors (counties) that are closer to the plants—with this design I can test whether this pattern holds

[14] I do not include NO_x and SO_2 emission rate here for two reasons. One, these values are undefined when monitors (counties) have no plants in a given distance bin. Two, the average value here may be somewhat uninformative as monitors (counties) with the same averages may have dramatically different patterns of underlying rates in their nearby plants.

[15] In both cases, the indicator is equal to 1 if any plant within 80 miles of the monitor (county centroid) was in a nonattainment zone or under the NO_x Budget Trading Program.

empirically. In both cases, there are no evident violations of the identifying assumptions. Air quality analyses are clustered at the monitor level,

3.4.2 Data and Sample Selection

Data for this project comes from several sources. First, data on annual unit-level generation, heat input and emissions for coal-fired power plants comes from the EPA's Continuous Emissions Monitoring System (CEMS) for all plants that were included in the Acid Rain Program, which includes almost all coal-fired plants in the contiguous United States. CEMS also includes data on plant characteristics including year of operation and latitude and longitude. My sample consists of the years 1998 to 2014.[16] Only the 1,068 coal-fired units with positive heat input in 1998, the baseline year, are included. Data on NAAQS county attainment status were taken from the EPA Green Book.

Second, data on New Source Review settlements comes from the Environmental Protection Agency's New Source Review Initiative website, which includes links to consent decrees of all settled cases. Using this information, I created a dataset of all plants affected by consent decrees, the date the case was initiated and the date the consent decree was signed. I matched each named plant to pollution data which could then be merged with CEMS data. These data were also cross-referenced with the EPA Enforcement and Compliance History Online (ECHO) database to identify non-settling cases, case milestone dates and penalty amounts.

Third, data on air pollution comes from the EPA air pollution monitoring system and data were obtained from AQS Datamart. Only monitors in use for all years between 1999 and 2014 are used to control for endogenous monitor openings/closings. Fourth, data on mortality comes from the Center for Disease Control (CDC) WONDER Database for years between 1998

[16]A new regulatory regime, the Cross-State Air Pollution Rule (CSAPR), came into effect in 2015, which put a new price on sulfur dioxide and may have changed plant decisionmaking. While this should not have affected settling and non-settling plants differentially conditional on covariates (and results are robust to including the full sample), my primary sample includes only years prior to 2014 to avoid the possibility that CSPAR may have influenced plant settlement and investment decisions. I show robustness to including all years in my sample.

and 2014. I use age-adjusted mortality rates and drop county-years with fewer than 20 deaths, which are deemed by the CDC to be “unreliable.” These rates typically occur in low-population counties and in order to prevent selection in my sample based on population or mortality changes, I restrict my sample to only counties with greater than 25,000 population in 1998—less than 3% of county-years in this group have rates that are “unreliable” for any outcome and less than 1% of county-years have unreliable cardiovascular mortality rates. I show robustness to relaxing this population restriction. Finally, data on utility revenues, pricing and ownership structure come from the Energy Information Administration (EIA) form 861.

3.5 Results

3.5.1 Pollution

Consent decrees typically included requirements to install SO_2 and NO_x control equipment, undertake other actions to reduce total SO_2 and NO_x pollution or decrease emission rates below a certain threshold. In this section, I estimate whether these requirements succeeded in reducing emissions. Figure 3.2 shows the primary results for total emissions and emission rates of SO_2 and NO_x pollution. Across all outcomes, there appear to be no major differences among settling and non-settling plants conditional on covariates in the lead up to settlement. However, pollution emissions begins to decrease in the year of a settlement and this decrease grows larger as time progresses. The size of the decline is quite large—by six years after a settlement, emissions have fallen by more than 15%. Further, this decline occurs for both total emissions and emission rates suggesting that it is not simply caused by changes in output.

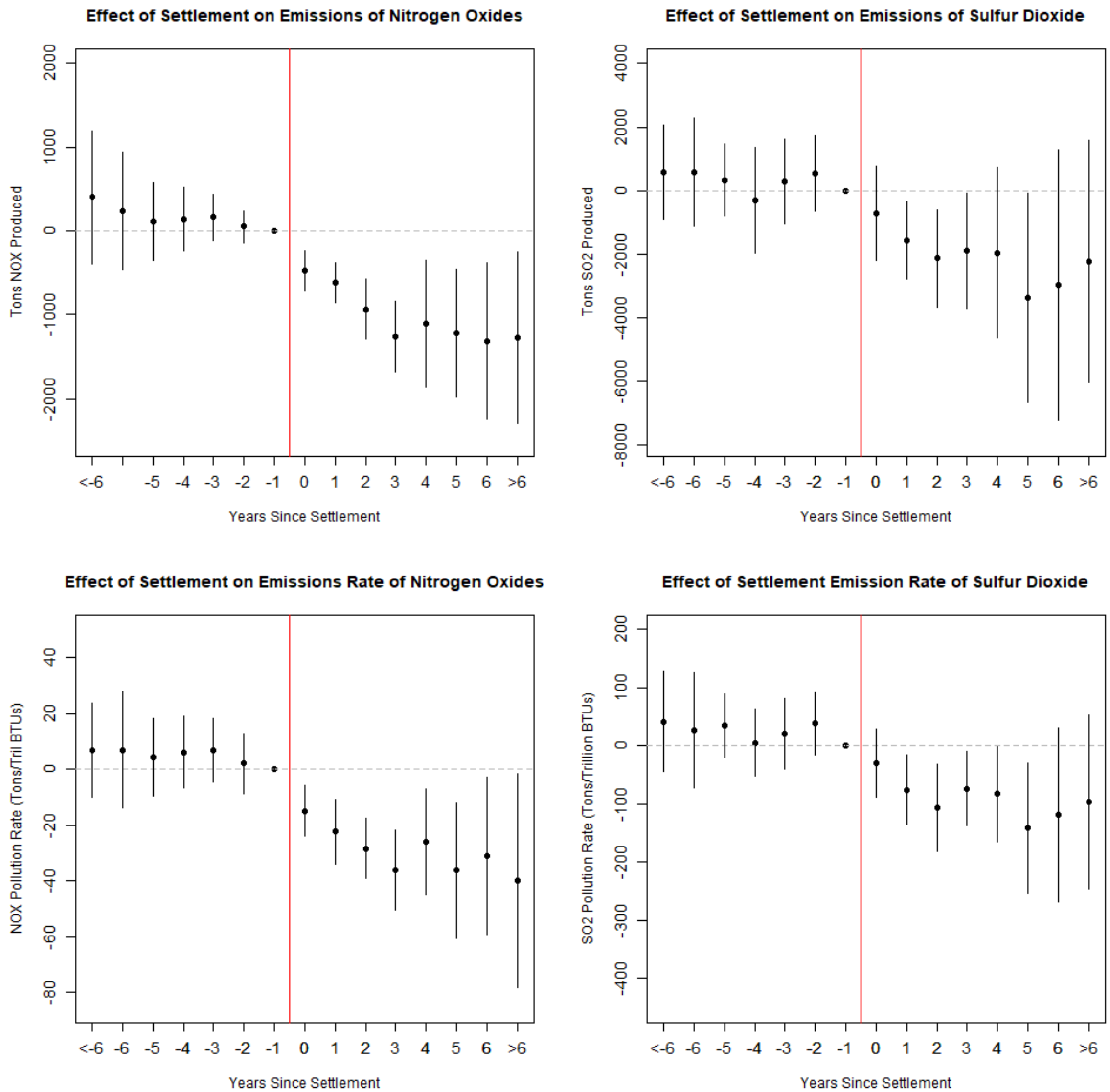


Figure 3.2: Effect of Settlement on Pollution and Output in Coal-Fired Power Plants

This figure shows the effect of a settlement on total emissions and emission rates among coal-fired power plants. Coefficients come from a unit-level regression of the outcome variable on indicators for years relative to settlement (never-settling plants have a value of 0 for all years), with -1 years to settlement the omitted variable. Never-settling units have a value of 0 for all indicators. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status by year fixed effects and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level.

Table 3.2: Effect of Settlement on Coal-Fired Power Plant Pollution and Generation Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NO_x	SO_2	NO_x rate	SO_2 rate	Heat Input (MMBTU)	Gross load	Retired	SO_2
Post Consent Decree	-1,069***	-2,105**	-30.99***	-77.94	0.274	0.000375	-0.0326	0.0435
	(189.3)	(861.0)	(6.432)	(55.71)	(0.547)	(0.0616)	(0.0221)	(0.0577)
Observations	17,421	17,421	14,972	14,972	17,421	17,421	17,421	17,421
R^2	0.863	0.856	0.878	0.886	0.969	0.969	0.742	0.968
Dep. Var Mean (1998)	5707	12366	276.3	733.4	21.91	2.069	0	2.240
Units	Tons	Tons	Tons/Tril. BTU	Tons/Tril. BTU	Tril BTU	TWh	Share	Mil. Short Tons

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. “Post Consent Decree” is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the “Post Consent Decree” variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a “Post Consent Decree” value of 0 for all years. The number of observations for emission rates shrinks because I drop years with less than 1 Tril BTU of heat input in a given year; rates become unstable with low-levels of heat input. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status by year fixed effects and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. Regressions are weighted by a unit’s share of a plant’s heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped.

Table 3.3: Effect of Settlement on Coal-Fired Power Plant Pollution and Generation Outcomes: By Litigation Status

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NO_x	SO_2	NO_x rate	SO_2 rate	Heat Input	Gross load	Retired	CO_2
Post Litig, Pre-Settlement	-91.05	-32.37	-4.276	15.30	-0.141	0.0297	0.0101	-0.0212
	(173.0)	(609.0)	(5.290)	(36.13)	(0.520)	(0.0481)	(0.0269)	(0.0525)
Post Consent Decree	-1,153***	-2,617***	-35.06***	-78.25*	0.179	-0.00465	-0.0232	0.0305
	(231.6)	(872.0)	(7.002)	(45.93)	(0.644)	(0.0629)	(0.0316)	(0.0683)
Observations	16,401	16,401	14,015	14,015	16,401	16,401	16,401	16,401
R^2	0.878	0.862	0.882	0.888	0.969	0.970	0.742	0.968
Dep. Var Mean (1998)	5493	12132	273.6	729.2	21.79	2.050	0	2.227
Units	Tons	Tons	Tons/Tril. BTU	Tons/Tril. BTU	Tril BTU	TWh	Share	Mil. Short Tons

Clustered standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

This table shows the effect of the onset of litigation and settlement on pollution and generation outcomes among coal-fired power plants. “Post Litigation, Pre-Settlement” is equal to 1 if any unit in a plant is under litigation, but has not yet settled in a given year. “Post Settlement” is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which litigation started or the consent decree was signed, these variables are equal to 1 if the event occurred prior to July of that year. Never-litigated units have a “Post Litigation, Pre-Settlement” and “Post Settlement” value of 0 for all years. Cases settled by the Tennessee Valley Authority are excluded as the original suit against the TVA was thrown out in the early 2000s, but the TVA then chose to settle in 2011, so it is unclear exactly when they perceived themselves to be “under litigation.” The number of observations for emission rates shrinks because I drop years with less than 1 million BTU of heat input; rates become unstable with very low-levels of heat input. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status by year fixed effects and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. Regressions are weighted by a unit’s share of a plant’s heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 966 units across 384 plants.

Table 3.2 shows the average treatment effect of coming under a consent decree. Consistent with Figure 3.2, the settlement led to large and significant declines in emissions. On average following a settlement, both SO_2 and NO_x emissions fell by 15%-20% relative to 1998 levels. There are also economically large declines in SO_2 and NO_x emission rates, although only NO_x rates are significant at conventional levels. The effect on total electricity generation (gross load) and heat input are small and insignificant implying that these changes happened almost entirely on the intensive margin. There is also no change in carbon-dioxide production suggesting that these settlements provided few climate-related co-benefits.[17]

Table 3.3 shows these effects with an additional indicator for whether or not a plant is in a post-litigation, but pre-settlement year.[18] This table shows that the effects of being in the pre-settlement, post-litigation period are very small and statistically insignificant, suggesting that litigated plants did not undergo large behavior change following the onset of the suit, at least relative to non-sued plants. Second, the effects of being in a post-settlement period relative to a pre-litigation period remain large and significant for all four pollution outcomes. This suggests the post-settlement results are not driven by responses to the onset of litigation, but instead represent declines even relative to the pre-litigation period. Further, the fact that I do not observe any significant differences between the pre-litigation and post-litigation, pre-settlement periods provides additional reassuring evidence that the observed effects are not merely driven by differential time trends among settling and non-settling plants.

Tables C.3, C.4, C.5 and C.6 attempt to address some of the identification concerns identified in Section 3. Each table shows robustness checks for a different primary outcome (NO_x emissions, NO_x emission rates, SO_2 emissions, and SO_2 emission rates respectively). Columns 1 and 2 show results are robust to constructing cells with 1998 generation and pollution quartiles (1) or cells additionally including 5-year bins for year of unit construction (2). Column (3) shows that results are also robust to including only state by year fixed effects—there is nothing in the cell-based specification that mechanically creates these results. Columns (4) and (5) shows that results are robust to only including settling plants. Column 5 removes cell by year fixed effects

[17] The EPA CEMS data does not track emissions of other pollutants, which may also be reduced as a byproduct of action taken to reduce SO_2 emissions. Thus, other co-beneifts from the policy may exist, but I am unable to examine them here.

[18] Plants affected by the TVA settlement are excluded from this analysis. The enforcement action against the TVA began in 1999 was deemed unconstitutional in 2001 by the 11th Circuit Court of Appeals and then never pursued. In 2011, the Obama administration created a new consent decree with the TVA based on the claims of this original action. It is therefore unclear whether (or when) the TVA perceived itself to be under litigation during the intervening period.

and state by year fixed effects as there is limited variation in settlement timing of plants within states, creating very large standard errors in Column (4), but in both cases point estimates are very similar. Column 6 shows effects using settlements that occur in groups of plants (>3)—results for NO_x remain largely unchanged, but effects for SO_2 do shrink and become statistically insignificant, although they remain highly economically significant. Column 7 shows results for all years (including 2015-2017 when CSPAR was in effect)—results do not change. Finally, Column 8 shows that results are robust to logging the dependent variable, suggesting that results are not driven by changes at the largest or highest polluting plants alone. Together these results add confidence that I am indeed estimating a causal effect of the settlements.

Table C.7 further tests robustness to exposure to federal and state cap and trade programs. In theory, it is possible that differential exposure to these programs even conditional on covariates could both influence settlement decisions and future pollution control investment decisions creating bias in my estimates. Two major regional cap-and-trade programs are the Regional Greenhouse Gas Initiative and independent SO_2 and NO_x emissions trading programs run by the state of Texas. In Columns (1) and (2), I exclude these states from the regression for the periods in which their cap-and-trade programs were in place and the results remain largely unchanged. An additional concern is that unit's may be differentially affected by exposure to federal emissions trading programs including the Acid Rain Program and the NO_x Budget Trading Program. The cell-based approach implicitly controls for many of the plant's baseline characteristics that we might expect to lead to differential impact across plants, but it is possible the cells are not granular enough to fully control for these differences. Accordingly, in Columns (3)-(6), I use even more granular controls for emission rates, which are likely to be the main determinant of differential response to these programs prices. Specifically in Columns (3)-(4), I control for 1998 NO_x SO_2 emission rate ventile by year fixed effects, while in columns (5)-(6), I control for state by 1998 emission rate ventile by year fixed effects. These restrictive controls ensure that I am comparing units that have very similar baseline pollution profiles, but reassuringly results remain very similar to those in the primary analysis.

I next show the effects of these settlements on the adoption of new pollution control equipment.[19] Tables 3.4 and 3.5 show the results for an analogous regression to the pollution and generation analysis, but looking at the adoption of NO_x and SO_2 control technology. In both cases, we see that settlements do not have large effects on the extensive margin of technology adoption. However, the settlement does appear to lead units to upgrade existing equipment. The settlement increases the probability that a new pollution control system is installed in a given year by 2 percentage points for NO_x and 1 percentage points for SO_2. Although the effects are not statistically significant at conventional levels, they represent a very large proportional effect. These point estimates imply a 40% increase in annual installation rate of new equipment for new NO_x control equipment relative to baseline and a more than doubling of the installation rate of new SO_2 control technology.

These newly-installed systems also appear to be different in type from existing technologies although again point estimates, while economically significant, are generally not statistically significant at conventional levels. Units are induced to switch from overfire air systems to selective catalytic and non-catalytic reduction systems, which provide higher levels of NO_x control. Similarly, plants are induced to increase their dry-scrubber SO_2 technology. There is also suggestive evidence that plants are adopting lower sulfur coal in response to settlements—the average sulfur content of burned coal falls by 8% in the six years after a settlement is in place, although again this effect is not statistically significant.[20]

[19]For retired years, I code each plant as having the pollution control equipment it had at retirement. For outcomes examining whether or not a plant installed a new piece of equipment in a given year or examining the sulfur content of a plant's coal, I examine pollution control adoption only in years in which units were operating. In both cases, it is therefore important to note that although there was no evidence of settlement causing large changes in retirement rates, if the types of units induced to retire changed in unobservable ways because of settlements, these results may be biased. Retirement rates are relatively low over the entire sample since it goes only until 2014 and so it is unlikely that this is driving the observed results. However, these results should nonetheless be interpreted with more caution.

[20]There is no effect in the absolute level of sulfur content, but I also observe an effect on the absolute level of sulfur content if several plants that use coal with very high sulfur content is excluded.

Table 3.4: Effects of Settlement on Coal-Fired Power Plants NO_x Pollution Control

	(1)	(2)	(3)	(4)	(5)	(6)
	Any NO_x Contr	New NO_x Contr/Yr	SCR Contr	OFA Contr	Low NO_x Contr	Oth Contr
Post Consent Decree	0.0259	0.0222	0.0662	-0.0312	0.0690*	0.00141
	(0.0614)	(0.0193)	(0.0510)	(0.0374)	(0.0395)	(0.00594)
Observations	17,421	15,853	17,421	17,421	17,421	17,421
R^2	0.868	0.346	0.838	0.892	0.896	0.893
Dep. Var Mean (1998)	0.527	0.0501	0.0172	0.145	0.445	0.0311

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on NO_x pollution control installation among coal-fired power plants. "Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. SCR refers to selected catalyic and non-catalytic reduction, OFA refers to overfire air, low NO_x controls refers to low NO_x burners and cells and other controls refer to all other controls. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped.

Table 3.5: Effects of Settlement on Coal-Fired Power Plants SO_2 Pollution Control

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Any SO_2 Contr	New SO_2 Contr/Yr	Wet Scrubber	Dry Scrubber	Oth Contr	Sulf Content	Ln Sulf Content
Post Consent Decree	0.00547	0.0104	-0.0158	0.0244	0.00189	-0.000481	-0.0836
	(0.0305)	(0.0148)	(0.0327)	(0.0197)	(0.00145)	(0.00111)	(0.0707)
Observations	17,421	15,853	17,421	17,421	17,421	13,822	14,347
R^2	0.898	0.362	0.907	0.880	0.934	0.939	0.952
Dep. Var Mean (1998)	0.221	0.00438	0.160	0.0310	0.0244	0.0121	-0.0946

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on SO_2 pollution control installation among coal-fired power plants. "Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. Sulfur content refers to a plant's coal input average sulfur content in a given year. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the plant level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped.

3.5.2 Air Quality

I next examine the effect of settlements on ambient air quality around affected plants. Examining the effect of the settlement on ambient air quality in addition to changes in pollution emissions from CEMS data is important for three reasons. First, although CEMS data are regularly audited, it may still be possible for affected plants to manipulate their data. Second, in theory, pollution decreases from affected plants may be offset by changes in behavior from other nearby power sources or industrial facilities. Third, ambient pollution levels are the relevant outcome for impacts on health and well-being.

I examine the effect of settlements on three major pollutants that have the potential to be affected by changes in coal-fired power plant pollution: SO_2, ozone, and PM 2.5.[21] SO_2 should be affected directly by the units' control strategies, while NO_x is a precursor to ozone and both NO_x and SO_2 are precursors to PM 2.5. Table 3.6 shows the main results. Odd columns show results controlling for cell by year fixed effects where cells are defined as all monitors in the same state, all above/below median heat input of plants within 40 miles and all above/below median of heat input of plants of plants 40-80 miles away, while even columns show specifications with cell by year by baseline (1998) outcome median. In Columns (1)-(2), we first see evidence for a large reduction in sulfur-dioxide pollution; having a plant come under consent decree within 40 miles of a monitor leads to a 3.5 ppb decline in SO_2 (15% decline). There is no effect for monitors 40-80 miles from a plant with a consent decree. This large proportional decline is exactly what would we expect given the observed effect of settlements on unit SO_2 emissions; coal-fired power plants are one of the largest sources of sulfur dioxide pollution in the United States and so plant-level reductions in sulfur dioxide should lead to large proportional reductions in ambient sulfur dioxide levels.

In Columns (3) and (4), we can see that consent decrees do not appear to lead to any economically or statistically significant changes in ozone levels. NO_x from coal-fired power plants are a relatively small contributor to overall ozone levels and ozone exhibits significant variation due to local weather patterns, so such a null result is perhaps unsurprising. Finally, we see that PM 2.5 declines both among monitors within 40 miles of a plant with a consent decree and monitors 40-80 miles from a plant with a consent decree, but effects are slightly larger and only significant for monitors 40-80 miles away. The reduction is relatively small (.75%-1.5% decline), consistent with coal-fired power plants again being a relatively small contributor to overall PM 2.5 levels. While we would expect the effect to be larger for monitors closer to the plants, the difference between the two coefficients is not statistically significant and tall smokestacks on many coal-fired power plants allow particles to travel many miles from the plant. Further, as we

[21] NO_2 should also be affected, but there are relatively few monitors within 40 miles of a plant so I lack sufficient power to examine this outcome.

will see in Table C.8, there is no effect for monitors 50-100 miles or 60-120 miles suggesting that these effects are concentrated among monitors that are plausibly close enough to the treated plants to be affected by their change in emissions.

Figure C.1 show the effect of a consent decree occurring within 40 miles of a monitor in event study form. Given the smaller sample, these estimates are noisier than our estimates of CEMS-based pollution. Nonetheless, there is little evidence of strong pre-trends for either of the three pollutants prior to a monitor's first settlement at a plant within 40 miles and SO_2 and to a lesser-extent, PM 2.5, exhibit clear and significant declines after this settlement comes into place. This provides further reassuring evidence that pollution declines are driven by plant settlements and not other correlated factors.

Table 3.6: Effect of Settlement on Ambient Air Quality

	(1)	(2)	(3)	(4)	(5)	(6)
	SO_2	SO_2	Ozone	Ozone	PM 2.5	PM 2.5
Post Consent Decree w/i 40 Mi	-3.587***	-3.956**	0.000284	0.000177	-0.136	-0.108
	(1.368)	(1.512)	(0.000315)	(0.000335)	(0.125)	(0.143)
Post Consent Decree w/i 40-80 Mi	-0.322	-0.408	-0.000182	-0.000236	-0.201**	-0.205**
	(1.086)	(1.020)	(0.000208)	(0.000219)	(0.0807)	(0.0886)
Observations	2,181	1,822	6,866	6,476	5,056	4,760
R^2	0.919	0.944	0.915	0.922	0.962	0.969
Cell x Year FE	Y	Y	Y	Y	Y	Y
Cell x Year x BL Val FE	N	Y	N	Y	N	Y
Dep. Var Mean (1998)	19.28	19.28	0.0582	0.0582	14.18	14.18
Monitors	137	115	405	381	337	318
Units	ppb	ppb	ppm	ppm	$\mu g/m^3$	$\mu g/m^3$

Clustered standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

This table shows the effect of a settlement on observed pollution levels in monitors nearby the plant. Only monitors that began operation prior to 1999 and were still in operation in 2017 were included. Ozone is measured in parts per million and is the annual maximum daily 1-hour value. Sulfur dioxide is measured in parts per billion and is the average annual hourly value. PM 2.5 is measured in $\mu g/m^3$ and is the average annual average daily value. Coefficients come from a regression of the outcome variable on indicators for whether or not a monitor has a plant under consent decree within 40 miles and and within 40-80 miles. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. “BL Val FE” refer to an indicator if a monitor was above or below the median value for the outcome variable in the monitor’s baseline year (typically 1998 for SO_2 and ozone monitors and 1999 for $PM2.5$ monitors). Fewer monitors and observations exist in the even columns because monitors that are singletons within a given cell by year are dropped and the more restrictive cell definitions in the even columns increase the number of singletons. Standard errors are clustered at the monitor level.

Table C.8 shows the same specification using different distance cut-offs. The results for

SO_2 are very similar when looking at plants within 30 miles of a monitor that settle relative to those that are 30=60 miles away. When I expand the distance bins in Column (4) and (7) to 50 and 60 miles relative to 50-100 miles and 60-120 miles respectively, results shrink suggesting that the SO_2 effects are driven by monitors relatively close to the plant. The results for ozone are small and insignificant regardless of the cut-off used. Finally, for PM 2.5 when using the <30 mile vs 30-60 mile cut-off settlements appear to cause about a 1% decrease in pollution levels in both distance bins although the effect is not statistically significant. In the <50 mi vs 50-100 mile and < 60 mile vs 60-120 mile specifications, PM 2.5 drops by 1.5-2% ($p<.01$) in the closer distance bin, with a smaller and statistically insignificant decrease for monitors 50-100 and 60-120 miles away. Taken together, these three specifications suggest that the PM 2.5 decreases are indeed caused by consent decrees and that the effects may persist further from the plant than SO_2.

3.5.3 Mortality

I now turn to examining whether the reductions in pollution observed above affected local mortality rates. I use the same specification as above but with counties as the unit of analysis. Distances from the county to nearby coal plants are calculated using a county's population-weighted centroid. Table 3.7 shows the main results. As above, odd columns show results controlling for cell by year fixed effects where cells are defined as all counties in the same state, all above/below median heat input of plants within 40 miles and all above/below median of heat input of plants of plants 40-80 miles away. Even columns show specifications with cell by year by baseline (1998) outcome mortality rate median. Under both specifications, having a plant come under consent decree within 40 miles of a county centroid leads to 6.5 fewer cardiovascular and respiratory deaths per 100,000 population decline per year, or a 1.5% decrease. This decrease is driven by both cardiovascular (1% decline) and respiratory (2.4%) deaths. There are no economically or statistically significant effects of having a plant come under consent decree 40 to 80 miles from a county centroid for any of the three outcomes. Additionally, in Columns

(7)-(8), we see that there is no negative effect of consent decree on externally-caused deaths such as motor vehicle accidents, overdoses, suicides and homicides; coefficients are positive, small and statistically insignificant. Together, these results add confidence that the declines in cardiovascular and respiratory mortality are caused by the consent decrees themselves and not other confounding factors.

Table 3.7: Effect of Settlement on Cardiovascular, Respiratory and External Mortality Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ttl Card+Resp	Ttl Card+Resp	Ttl Card	Ttl Card	Ttl Resp	Ttl Resp	Ttl External	Ttl External
	Mort. Rate	Mort. Rate	Mort. Rate	Mort. Rate	Mort. Rate	Mort. Rate	Mort. Rate	Mort. Rate
Post Consent Decree w/i 40 Mi	-6.352**	-6.029***	-3.991**	-4.034**	-2.394**	-2.085**	0.104	0.176
	(2.507)	(2.196)	(1.863)	(1.762)	(1.066)	(0.840)	(0.852)	(0.905)
Post Consent Decree w/i 40-80 Mi	0.303	0.789	0.143	0.792	0.168	0.190	0.686	0.480
	(1.792)	(2.057)	(1.521)	(1.965)	(0.955)	(0.806)	(0.557)	(0.614)
Observations	21,470	20,705	21,493	20,754	21,354	20,072	19,597	17,359
R^2	0.893	0.902	0.888	0.897	0.679	0.721	0.775	0.799
Cell x Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Extended Cell x Year FE	N	Y	N	Y	N	Y	N	Y
Dep. Var. Mean (1998)	468.5	468.5	373.1	373.1	95.75	95.75	62.72	62.72
Units	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Units	/100K Pop	/100K Pop	/100K Pop	/100K Pop	/100K Pop	/100K Pop	/100K Pop	/100K Pop
Counties	1265	1219	1265	1221	1264	1195	1261	1062

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of a coal-fired power plant are included. All mortality rates are age-adjusted and expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder because they are based on fewer than 20 deaths in a year are excluded. Only counties with greater than 25,000 population in 1998 are included. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Counties in the same extended cells must also be all above/below median of the outcome variable at baseline. Fewer counties and observations exist in the even columns because monitors that are singletons within a given cell by year are dropped and the more restrictive cell definitions in the even columns increase the number of singletons. There are more observations for the cardiovascular mortality rate outcome because cardiovascular deaths are more common than respiratory deaths, so more counties have reliable cardiovascular death rates than respiratory death rates. There are fewer cardiovascular+respiratory mortality observations than cardiovascular mortality alone because several counties with data for cardiovascular mortality have suppressed data on respiratory mortality (fewer than 10 deaths in a year), making cardiovascular+respiratory mortality undefined. Standard errors are clustered at the state level.

Figure C.2 shows the dynamic effect of a consent decree occurring within 40 miles of a county centroid in order to test whether the observed results could be driven by differential trends among treated counties.[22] As with the air quality analysis, the individual year coefficients are noisy but there appears to be no trends prior to the consent decree for either cardiovascular or respiratory deaths. Following the consent decree, both cardiovascular and respiratory deaths exhibit a steady decline in mortality rate.

Table C.9 shows the results of the primary analysis without using the 25,000 population cut-off. Effects for cardiovascular+respiratory mortality and respiratory mortality alone remain statistically and economically significant. The coefficient size for cardiovascular mortality shrinks by about a third and is no longer statistically significant, but remains highly economically significant (2.5 deaths per 100,000 population per year reduction). Table C.10 shows results when using the crude death rate instead of an age-adjusted death rate. Results again remain largely unchanged and if anything increase. Table C.11 shows results when using different distance cut-offs. In Columns (1)-(3), we can see that both counties 0-30 miles and 30-60 miles away from a settling plant appear to have meaningful mortality reductions for cardiovascular mortality, although the effect in both bins is no longer statistically significant. The effect for respiratory mortality remains highly statistically significant and is concentrated in counties whose centroids are less than 30 miles from the treated unit. In Columns (4)-(9), we see that results for the < 50 mile and <60 mile are similar, but slightly smaller than the coefficients observed in the primary specifications. As in the main specification, there is no effect on mortality when a plant comes under consent decree 50-100 miles or 60-120 miles from a county centroid.

The specifications above estimate the effect of having at least one plant within a given distance of a county centroid under consent decree. This makes the total effect somewhat difficult to interpret—some counties have multiple plants with consent decrees within a given distance bin and we are averaging effects across all such counties. Table C.12 attempts to address this concern by showing effects separately for counties that have a single nearby plant under consent decree and counties that have multiple nearby plants come under consent decree. There are two important

[22] All estimates come from regressions that continue to control for consent decrees 40-80 miles away as in the main specification.

takeaways from this table. First, the effect of having a single nearby plant under consent decree on cardiovascular and respiratory mortality remains economically large and statistically significant. Second, effect sizes are larger for counties with multiple plants under consent decree—this is precisely the pattern we would expect if the mortality decline was truly caused by the changes induced by the consent decree and not other confounding factors.

The results from both the CEMS and air monitor analysis suggest that the New Source Review litigation was extremely successful in reducing air pollution from affected plants. I show further suggestive evidence that these pollution declines led to decreases in mortality—evidence that is bolstered by much previous research establishing links between sulfur-dioxide pollution and health outcomes (Luechinger, 2014; Barreca, Neidell, and Sanders, 2017). However, these pollution improvements also came with large costs for utilities. Accordingly, in the next section, I discuss the incidence of the costs imposed by litigation, examining whether these costs were borne by ratepayers or other utility stakeholders.

3.5.4 Settlement Incidence

New Source Review litigation settlements created substantial costs for settling firms. The EPA estimated that total compliance costs were more than $20 billion across all firms.[23] Settlements also likely created significant indirect costs of compliance not included in the EPA's estimates such as the increased expense of power generation once pollution control equipment was installed. In this section, I analyze the extent to which these costs were passed onto consumers through increases in retail electricity prices. To do this, I first restrict my sample to utilities that both operated plants and distributed electricity—this excludes utilities in states with restructured electricity markets that only sell electricity on the wholesale market. The expected effect of settlements on electricity prices (and firm revenues) in restructured markets will depend upon

[23] Author's calculation from EPA ECHO database. The EPA's estimates are likely highly imprecise. However, even if overestimated by an order of magnitude, these compliance costs would still be quite large.

where affected plants were in the dispatch order prior to the settlement and how settlement shifted the costs of the region's marginal unit. I lack sufficient variation to provide good empirical estimates of these changes and so I focus only on the effect of the policy in regulated markets.

In regulated markets, we might expect that utilities will attempt to pass on the increased capital costs arising from settlements onto ratepayers. However, such rate increases must be approved by state Public Utility Commissions and it is ex-ante unclear the extent to which these commissions will allow litigation-related expenses to be passed onto local consumers. In this section, I test for such pass through empirically by comparing changes in average electricity prices and utility revenues before and after a settlement relative to other utilities that either had not yet settled or were never under litigation. Because 89% of settling plants were large, investor-owned utilities, I focus my analysis on only utilities that meet the following criteria in order to create comparable counterfactuals: privately-owned utilities with a 1998 customer base greater than 25,000. I show robustness to relaxing both of these constraints. This leaves me with a sample of 123 utilities of which 20 had at least one plant come under a consent decree.

I use a similar cell-based empirical strategy as in the analyses above. I create cells that attempt to group together utilities that are similar in terms of size, customer base and energy generation mix. Specifically, cells consist of utilities that are above/below median 1998 customers, above/below median 1998 average electricity price, above/below median average share of generation from coal and above/below median average share of generation generated by natural gas. Because there are only 123 utilities, some of which operate in multiple states, I do not include state by cell by year fixed effects in main analysis. However, I show robustness to North American Electric Reliability Corporation (NERC) region by cell by year fixed effects for all outcome variables. Standard errors are clustered at the utility level.

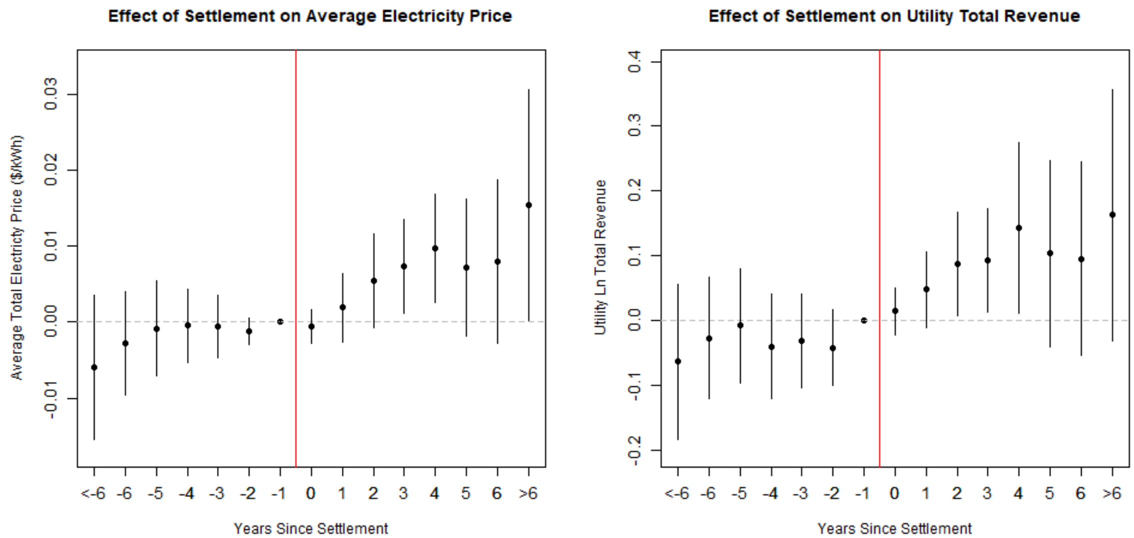


Figure 3.3: Effect of Settlement on Utility Prices and Revenue

This figure shows the effect of a settlement on utility average retail price and log total revenue. Only investor-owned utilities that both operate plants and distribute electricity and had greater than 25,000 population in 1998 are included. Settlement is considered the year of the utility's first New Source Review settlement. Coefficients come from a regression of the outcome variable on indicators for years relative to settlement (never-settling utilities have a value of 0 for all years). All regressions include utility fixed effects and cell by year fixed effects. Cells consist of all utilities that are both above/below median 1998 customer level ,above/below median 1998 average price, above/below median 1998 coal share of generation and above/below median 1998 natural gas share of generation. All standard errors are clustered at the utility level.

Table 3.8: Effect of Settlement on Utility Price, Revenue and Usage Outcomes

	(1) Average Price	(2) Average Price	(3) Total Revenue	(4) Total Revenue	(5) Total Customers	(6) Total Customers	(7) Annual Usage/ Customer	(8) Annual Usage/ Customer
Post Consent Decree	0.00949***	0.00523	0.129**	0.145*	0.0253	0.0162	-0.481	0.865
	(0.00353)	(0.00519)	(0.0530)	(0.0839)	(0.0324)	(0.0391)	(0.818)	(0.870)
Observations	1,849	1,322	1,849	1,231	1,849	1,322	1,849	1,322
R^2	0.877	0.914	0.969	0.989	0.982	0.988	0.946	0.969
Utility FE	Y	Y	Y	Y	Y	Y	Y	Y
Cell x Year FE	Y	Y	Y	Y	Y	Y	Y	Y
NERC x Cell x Year FE	N	Y	N	Y	N	Y	N	Y
Dep. Var Mean (1998)	0.106	0.106	1.778e+06	1.778e+06	669507	669507	26.84	26.84
Units	$/kWh	$/kWh	’000 $	’000 $	Cust.	Cust	MWh/Cust	MWh/Cust
Utilities	122	91	122	91	122	91	122	91

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on utility average retail price, log total revenue, log total customers and average energy use per customer. Only investor-owned utilities that both operate plants and distribute electricity and had greater than 25,000 population in 1998 are included. Settlement is considered the year of the utility’s first New Source Review settlement. Coefficients come from a regression of the outcome variable on an indicator for whether or not a utility has signed a consent decree (never-settlers have a 0 for all years). All regressions include utility fixed effects and cell by year fixed effects. Cells consist of all utilities that are both above/below median 1998 customer level ,above/below median 1998 average price, above/below median 1998 coal share of generation and above/below median 1998 natural gas share of generation. In even columns, utilities in the same cell must also be in the same NERC region. Regressions without NERC fixed effects contain 122 utilities, while regressions with NERC fixed effects have 91 utilites. These differences from the 123 utilities that meet the sample criteria occur because utilities that are the only observation within their cells are dropped as cell-by-year fixed-effects become unidentified—more singletons exist as cells become more granular with the inclusion of NERC by year fixed effects. All dependent variable means are unlogged for logged outcome variables. All standard errors are clustered at the utility level.

Figure 3.3 show the main results for average retail prices and log utility revenue. Average retail price is defined as total annual utility retail revenue divided by total annual utility electricity sold.[24] There appears to be little trend in price or total revenue per customer prior to the settlement.

[24]Many utilities use block pricing schemes—unfortunately the best dataset on the levels of such block rates begins

After the settlement, both prices and revenues increase dramatically—by 4 or 5 years after the settlement both prices and revenues have increased by approximately 10%-15% and the difference relative to the pre-settlement year is statistically significant. Table 3.8 shows the average treatment effect of this specification both with (even columns) and without (odd columns) NERC by year fixed effects. The settlement led to a \$.009 increase in price per kWh and a 14% increase in total revenue following the settlement. Effects are less precise when I add in NERC by cell by year fixed effects, but remain broadly similar. Reassuringly, there is little evidence for large changes in the total number of customers after a consent decree. There is also little evidence for changes in customers' average annual electricity use.[25]

In Tables C.13 and C.14, I show that these results are not reliant on specific specification choices or sample inclusion criteria for either average price (Table C.13) or total revenue (Table C.14). In Column (1) I show that results are robust to using both levels and logs. In Column (2) I show results are unchanged if we restrict our sample to only those utilities with data for all years between 1999 and 2014. In Column (3), I show results are similar if we only look at six years before and after the settlement date. In Columns (4)-(5), I show results are robust to fewer fixed effects (NERC by year only) or more restrictive fixed effects (energy source share quartiles). In Columns (6), I show results are unchanged if we restrict our sample to pre-2011 when the natural gas boom accelerated, which may have differentially affected electricity pricing. Finally, in Columns (7)-(9), I show results are broadly similar to the main specification when including IOUs with fewer than 25,000 customers (Column 7), including publicly-owned utilities with greater than 25,000 population (Column 8) and including all utilities that both generate and transmit electricity (Column 9), suggesting that the sample exclusion criteria are not driving the results.

These results provide suggestive evidence that, at least among regulated IOUs, a sub-

only in 2014 (Borenstein and Bushnell, 2018) and so I lack data on such rates for almost my entire sample. As a result, I instead use the average rate paid, which I can calculate for all years that a utility reported data to the EIA.

[25] Although, we may expect use to decline in response to average prices, I am under-powered to detect small changes and so would not expect to be able to detect such changes here.

stantial proportion of the costs of environmental improvements from litigation are passed onto ratepayers. Results from Table C.14 suggest that the average settling utility had an extra 197 million dollars per year after the settlement. There were twenty settling IOUs in states with regulated electricity markets and cumulative compliance costs as reported by the EPA among these utilities came to roughly 15 billion dollars. Thus, a back-of-an-envelope calculation suggests that under reasonable discount rate assumptions, almost the entire compliance cost as reported by the EPA was passed onto ratepayers rather than other utility stakeholders. However it is important to note that there are likely additional costs that the EPA did not estimate including foregone revenue from an increase in the cost of producing electricity, litigation costs and the opportunity cost of using $15 billion dollars in compliance investments. Further, the accuracy of the EPA estimates of compliance costs is unknown, although the EPA likely had an incentive to overestimate rather than underestimate costs. As a result, the true share of costs created by the litigation that were passed onto ratepayers remains highly uncertain.

3.6 Conclusion

Litigation is an important tool available to policymakers when rules and regulations are written ambiguously. This is particularly true in the environmental setting, which is governed by thousands of pages of complex administrative law. However, even if litigation is successful in extracting settlements from sued firms, there are a number of reasons to believe that these settlements will not lead to the desired policy outcomes. In particular, settlements may include requirements that the litigated firm would have complied with even in the absence of litigation. Further, any settlement agreements must be enforced, which may be difficult if the political and policy priorities of the government change. Finally, the incidence of costs imposed by the settlement is unclear.

In this paper, I provide new evidence on these questions in the context of the coal-fired power plant initiative, a EPA and Department of Justice program that sued coal-fired power plants for violations of the New Source Review provision of the Clean Air Act. The initiative led to more than $21 billion in compliance activities as estimated by the EPA and led one-third of US coal-fired power plants to come under consent decree. I show that these consent decrees did indeed lead to large (15-20%) reductions in pollution emissions. This reduction led to improvements in local ambient air quality and decreases in local cardiovascular and respiratory mortality. I further show suggestive evidence that among regulated utilities, a significant share of the costs for these large reductions were borne by ratepayers. In sum, these results suggest that litigation initiatives have the potential to be a successful policy tool in the environmental setting, but that any gains achieved in regulated utility markets may come at least partially at the expense of ratepayers rather than other utility stakeholders.

Chapter 3, in full, is currently under submission for publication of the material. Krumholz;Samuel. "The Effectiveness of Litigation as a Policy Instrument: The Case of the New Source Review Litigation". The dissertation author was the primary investigator and author of this material.

C.1 Tables and Figures

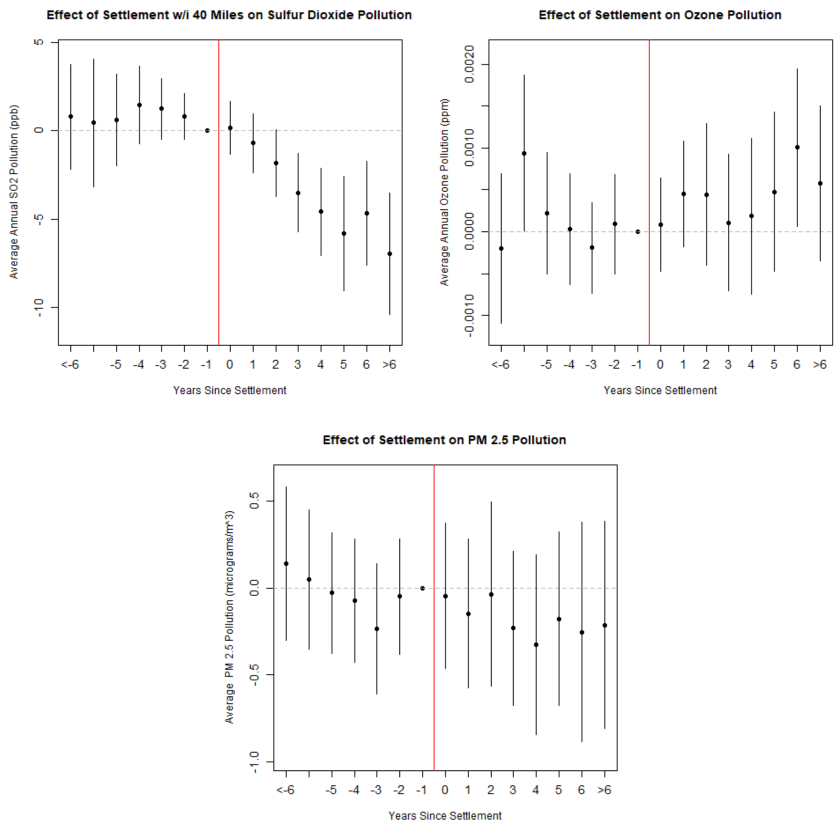


Figure C.1: Effect of Settlement on Ambient Air Quality Near Coal-Fired Power Plants

These figures show the effect of a settlement on observed pollution levels in monitors nearby the plant. Only monitors that began operation prior to 1999 (2000 for PM 2.5) and were still in operation in 2017 were included. Ozone is measured in parts per million and is the annual maximum daily 1-hour value. Sulfur dioxide is measured in parts per billion and is the average annual hourly value. PM 2.5 is measured in $\mu g/m^3$ and is the average annual average daily value. Coefficients come from a regression of the outcome variable on indicators for whether or not a monitor has a plant under consent decree within 40 miles and and within 40-80 miles. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all monitors in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Standard errors are clustered at the monitor level.

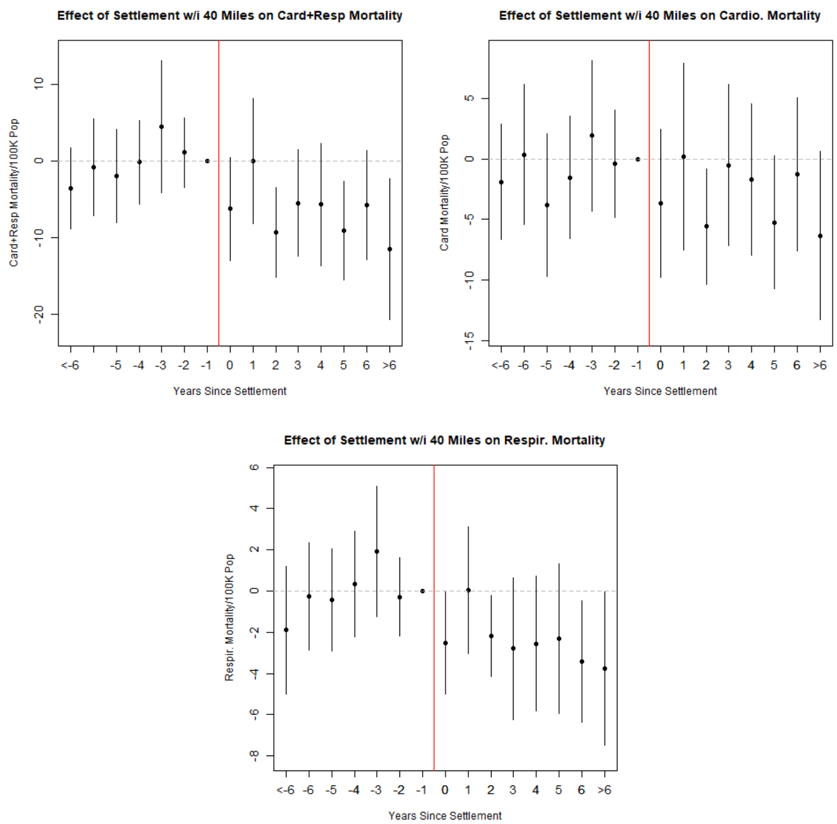


Figure C.2: Effect of Settlement on Mortality Rate in Counties Nearby Coal-Fired Power Plants

This figure shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of a coal-fired power plant are included. All mortality rates are age-adjusted and expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder are excluded. Only counties with greater than 25,000 population in 1998 are included. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Standard errors are clustered at the state level.

Table C.1: Summary Statistics of Primary Outcome Variables

Variable	N	Mean	SD	Min	Max
NO_x tons emitted	18,156	2,888	3,586	0	70,213
SO_2 tons emitted	18,156	7,343	10,0096	0	145,724
NO_x emission rate (tons/Tril BTU)	16,525	190	108	0	2,272
SO_2 emission rate (tons/Tril BTU)	16,525	576	531	0	16,679
Ambient SO_2 (ppb)	3,185	11.45	10.89	0	94.11
Ambient ozone (ppm)	8,594	.0514	.0062	.0273	.0762
Ambient $PM2.5$ (μ/m^3)	6,615	11.2	3.19	2.1	24.6
Age Adj Cardio Mort Rate (/100K Pop)	24,403	293	73	75	646
Age Adj Resp Mort Rate (/100K Pop)	23,967	88	23	30	261
Average Retail Elec. Price ($/kWh)	2,179	.099	.042	.037	.442
Average Utility Revenue (Mil $)	2,193	1,981	2,307	0	14,022

This table shows summary statistics for the primary outcome variables used in this paper. Mortality outcomes are only shown for counties with population greater than 25,000 that have reliable age-adjusted death rates as determined by the CDC. Utility prices and revenue are only shown for Investor Owned Utilities that both generate and transmit electricity and serve more than 25,000 customers. Both sets of restrictions are put in place to match the restrictions used in my primary analyses.

Table C.2: Conditional Correlation Between Treatment and Key Baseline Variables: State Fixed Effects

	(1) Heat input	(2) NO_x Rate	(3) SO_2 Rate	(4) Op Year	(5) Heat input	(6) NO_x Rate	(7) SO_2 Rate	(8) Op Year
Ever Settle	6.813***	3.879	17.62	0.129				
	(2.010)	(16.03)	(112.7)	(1.569)				
Settlement Year					-0.230	11.30**	4.168	-0.336
					(0.414)	(4.963)	(23.57)	(0.388)
Observations	1,065	1,065	1,065	1,065	350	350	350	350
R^2	0.194	0.157	0.221	0.267	0.155	0.196	0.315	0.190
Dep. Var Mean (1998)	18.89	289.4	794.3	1965	18.89	289.4	794.3	1965
Units	Tril. BTU	Tons/Tril BTU	Tons/Tril BTU	Years	Tril. BTU	Tons/Tril BTU	Tons/Tril BTU	Years

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of ever settling or settlement year on various baseline outcomes. All regressions include state fixed effects. There are 1,068 units included in the regression from 409 plants, but three units are the only units within their respective states and so are dropped. In Columns (5)-(8) only units that ever settle are included—there are 350 of such units from 109 plants. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year in order to not overweight plants with many units.

Table C.3: Effect of Settlement on Coal-Fired Power Plant NO_x Pollution: Robustness Check

	(1) NO_x	(2) NO_x	(3) NO_x	(4) NO_x	(5) NO_x	(6) NO_x	(7) NO_x	(8) Ln NO_x
Post Consent Decree	-1,357***	-1,327***	-1,707***	-720.4**	-1,018***	-1,149***	-968.5***	-0.326***
	(422.7)	(482.6)	(282.9)	(356.7)	(297.1)	(362.3)	(203.7)	(0.0416)
Observations	13,867	13,580	18,028	5,678	5,950	15,347	20,499	15,320
R^2	0.948	0.928	0.759	0.852	0.680	0.863	0.856	0.902
Spec	Cells w/ Quartiles	Cells w/ Open Yr	State x Year FE Only	Only Settling Plants	Only Settling Plants	>3 Plants in Stmnt	All Yrs	Logged Dep Var
Dep. Var Mean (1998)	5707	5707	5707	7309	7309	5667	5707	5707
Units	Tons	Tons	Tons	Tons	Tons	Tons	Tons	Tons

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. Never-settling units have a "post" value of 0 for all years. 'Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects unless otherwise stated. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. In Column (1) quartiles are used instead of medians. In Column (2) cells include year of operation bins (5 yr). In Column (3), cells are defined as state's only. Column (4) includes only settling plants with cell by year fixed effects, while Column (5) contains only settling plants with year fixed effects and unit fixed effects. Column (6) is restricted to settlements with more than 3 plants, Column (7) shows the base specification for all years between 1998 and 2018 and in Column (8), the dependent variable is logged. All dependent variable means are presented in levels. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped. Specifications with settling plants only contain 350 units at 109 plants. These estimates are clustered at the plant level, as only 20 states have a plant with a consent decree.

Table C.4: Effect of Settlement on Coal-Fired Power Plant NO_x Pollution Rate: Robustness Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NO_x rate	NO_x rate	NO_x rate	NO_x rate	NO_x rate	NO_x rate	NO_x rate	Ln NO_x rate
PPost Consent Decree	-46.36***	-33.49**	-37.30***	-26.25**	-26.16***	-26.77***	-31.14***	-0.308***
	(12.56)	(12.43)	(10.30)	(10.39)	(7.359)	(7.376)	(5.945)	(0.0416)
Observations	11,626	11,404	15,602	4,967	5,256	13,198	16,580	14,849
R^2	0.945	0.929	0.783	0.871	0.715	0.880	0.879	0.869
Spec	Cells w/ Quartiles	Cells w/ Open Yr	State x Year FE Only	Only Settling Plants	Only Settling Plants	>3 Plants in Stmnt	All Yrs	Logged Dep Var
Dep. Var Mean (1998)	275.1	275.1	275.1	307.7	307.7	272.9	275.1	275.1
Units	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. Never-settling units have a "post" value of 0 for all years. 'Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects unless otherwise stated. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. In Column (1) quartiles are used instead of medians. In Column (2) cells include year of operation bins (5 yr). In Column (3), cells are defined as state's only. Column (4) includes only settling plants with cell by year fixed effects, while Column (5) contains only settling plants with year fixed effects and unit fixed effects. Column (6) is restricted to settlements with more than 3 plants, Column (7) shows the base specification for all years between 1998 and 2018 and in Column (8), the dependent variable is logged. All dependent variable means are presented in levels. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped. Specifications with settling plants only contain 350 units at 109 plants. These estimates are clustered at the plant level, as only 20 states have a plant with a consent decree.

Table C.5: Effect of Settlement on Coal-Fired Power Plant SO_2 Pollution: Robustness Check

	(1) SO_2	(2) SO_2	(3) SO_2	(4) SO_2	(5) SO_2	(6) SO_2	(7) SO_2	(8) Ln SO_2
Post Consent Decree	-3,930**	-3,389**	-3,818***	-2,128	-2,790***	-876.0	-1,984**	-0.295**
	(1,697)	(1,391)	(1,160)	(2,275)	(864.6)	(1,475)	(756.8)	(0.141)
Observations	13,867	13,580	18,028	5,678	5,950	15,347	20,499	15,324
R^2	0.929	0.910	0.761	0.825	0.643	0.860	0.846	0.835
Spec	Cells w/ Quartiles	Cells w/ Open Yr	State x Year FE Only	Only Settling Plants	Only Settling Plants	>3 Plants in Stmnt	All Yrs	Logged Dep Var
Dep. Var Mean (1998)	12366	12366	12366	15696	15696	12415	12366	12366
Units	Tons	Tons	Tons	Tons	Tons	Tons	Tons	Tons

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. Never-settling units have a "post" value of 0 for all years. 'Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status unless otherwise stated and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. In Column (1) quartiles are used instead of medians. In Column (2) cells include year of operation bins (5 yr). In Column (3), cells are defined as state's only. Column (4) includes only settling plants with cell by year fixed effects, while Column (5) contains only settling plants with year fixed effects and unit fixed effects. Column (6) is restricted to settlements with more than 3 plants, Column (7) shows the base specification for all years between 1998 and 2018 and in Column (8), the dependent variable is logged. All dependent variable means are presented in levels. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped. Specifications with settling plants only contain 350 units at 109 plants. These estimates are clustered at the plant level, as only 20 states have a plant with a consent decree.

Table C.6: Effect of Settlement on Coal-Fired Power Plant SO_2 Pollution Rate: Robustness Check

	(1) SO_2 rate	(2) SO_2 rate	(3) SO_2 rate	(4) SO_2 rate	(5) SO_2 rate	(6) SO_2 rate	(7) SO_2 rate	(8) Ln SO_2 rate
Post Consent Decree	-122.2	-106.5	-103.7*	-134.8**	-101.2***	-42.38	-69.74	-0.300**
	(107.6)	(84.86)	(55.21)	(55.35)	(34.38)	(77.19)	(48.21)	(0.139)
Observations	11,626	11,404	15,602	4,967	5,256	13,198	16,580	14,850
R^2	0.952	0.939	0.821	0.902	0.762	0.890	0.888	0.836
Spec	Cells w/ Quartiles	Cells w/ Open Yr	State x Year FE Only	Only Settling Plants	Only Settling Plants	>3 Plants in Stmnt	All Yrs	Logged Dep Var
Dep. Var Mean (1998)	720.4	720.4	720.4	828.3	828.3	725	720.4	720.4
Units	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU	Tons/Tril BTU

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. Never-settling units have a "post" value of 0 for all years. 'Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. The number of observations for emission rates shrinks because I drop years with less than 1 MM BTU of heat input; rates become unstable with very low-levels of heat input. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status unless otherwise stated and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate. In Column (1) quartiles are used instead of medians. In Column (2) cells include year of operation bins (5 yr). In Column (3), cells are defined as state's only. Column (4) includes only settling plants with cell by year fixed effects, while Column (5) contains only settling plants with year fixed effects and unit fixed effects. Column (6) is restricted to settlements with more than 3 plants, Column (7) shows the base specification for all years between 1998 and 2018 and in Column (8), the dependent variable is logged. All dependent variable means are presented in levels. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level. These regressions includes 1,026 units across 395 plants—this is fewer than the 1,068 units in the full sample because all cells (including the NAAQS nonattainment indicators by year cells) that have only singleton values are dropped. Specifications with settling plants only contain 350 units at 109 plants. These estimates are clustered at the plant level, as only 20 states have a plant with a consent decree.

Table C.7: Effect of Settlement on Coal-Fired Power Plant Pollution and Generation Outcomes: Robustness to Sensitivity to Existing Cap-and-Trade Programs

	(1)	(2)	(3)	(4)	(5)	(6)
	NO_x Tons	SO_2 Tons	NO_x Tons	SO_2 Tons	NO_x Tons	SO_2 Tons
Post Consent Decree	-1,073***	-2,091**	-1,048***	-2,131**	-1,387***	-3,854**
	(191.1)	(849.8)	(184.9)	(852.8)	(248.4)	(1,480)
Observations	16,441	16,441	17,421	17,421	12,795	12,795
R^2	0.863	0.855	0.887	0.874	0.964	0.961
Spec	No RGGI States or TX	No RGGI States or TX	Emis. Ventiles x year	Emis. Ventiles x year	State x Emis. Ventiles x year	State x Emis. Ventiles x year

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. "Post Consent Decree" is equal to 1 if a year is after any unit in a plant came under a consent decree. For the year in which the consent decree was signed, the "Post Consent Decree" variable is equal to 1 if the consent decree happened prior to July of that year. Never-settling units have a "Post Consent Decree" value of 0 for all years. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status by year fixed effects and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate.NO RGGI states or Texas excludes all states in New England, New York, New Jersey, Delaware and Maryland from 2009-2014 and Texas for all years. "Emis. Ventiles" specifications include ventiles of 1998 NO_x and SO_2 emission rates interacted with year fixed effects. "State by Emis. Ventiles" include state by 1998 NO_x and SO_2 emission rate ventiles interacted with year. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level.

Table C.8: Effect of Settlement on Ambient Pollution Values: By Different Distance Cut-Offs

	(1) SO_2	(2) Ozone	(3) PM 2.5	(4) SO_2	(5) Ozone	(6) PM 2.5	(7) SO_2	(8) Ozone	(9) PM 2.5
Post Consent Decree w/i 30 Mi	-4.191***	0.000183	-0.178						
	(1.374)	(0.000320)	(0.127)						
Post Consent Decree w/i 30-60 Mi	0.0751	-0.000344	-0.130						
	(0.998)	(0.000267)	(0.127)						
Post Consent Decree w/i 50 Mi				-1.534	0.000216	-0.229*			
				(1.143)	(0.000311)	(0.121)			
Post Consent Decree w/i 50-100 Mi				-0.131	-0.000264	-0.0954			
				(1.071)	(0.000206)	(0.0768)			
Post Consent Decree w/i 60 Mi							-1.707	-0.000115	-0.288***
							(1.056)	(0.000283)	(0.0947)
Post Consent Decree w/i 60-120 Mi							-0.0669	-4.72e-05	-0.102
							(0.547)	(0.000197)	(0.0846)
Observations	2,033	6,348	4,589	2,240	7,283	5,298	2,512	7,810	5,506
R^2	0.928	0.909	0.961	0.912	0.913	0.962	0.924	0.911	0.962
Cell by Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dep. Var Mean (1998)	20	0.0587	14.51	19.14	0.0579	14.16	18.98	0.0577	14.11
Units	ppb	ppm	μ/m^3	ppb	ppm	μ/m^3	ppb	ppm	μ/m^3

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on observed pollution levels in monitors nearby the plant. Only monitors that began operation prior to 1999 (2000 for $PM2.5$, were still in operation in 2017 were included and are within 60 miles (cols (1)-(3)), 100 miles (Cols (4)-(6), or 120 miles (Cols (7)-(9)) of an ever settling plant are included. Only monitors that began operation prior to 1999 (2000 for PM 2.5) and were still in operation in 2017 were included. Ozone is measured in parts per million and is the annual maximum daily 1-hour value. Sulfur dioxide is measured in parts per billion and is the average annual hourly value. PM 2.5 is measured in $\mu g/m^3$ and is the average annual average daily value. Coefficients come from a regression of the outcome variable on indicators for whether or not a monitor has a plant under consent decree within 30,50, or 60 miles and and within 30-60, 50-100 or 60-120 miles. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 30,50 or 60 miles and below/above median 1998 heat input within 30-60,50-100 or 60-120 miles. All standard errors are clustered at the monitor level.

Table C.9: Effect of Settlement on Cardiovascular, Respiratory and Other Mortality: Full Sample

VARIABLES	(1) Ttl Card+Resp Mort Rate	(2) Ttl Card+Resp Mort Rate	(3) Ttl Card Mort Rate	(4) Ttl Card Mort Rate	(5) Ttl Resp Mort Rate	(6) Ttl Resp Mort Rate
Post Consent Decree w/i 40 Mi	-3.851*	-3.145*	-2.384	-2.490	-2.594***	-2.435**
	(2.172)	(1.751)	(1.713)	(1.536)	(0.920)	(0.987)
Post Consent Decree w/i 40-80 Mi	0.991	0.952	1.371	1.928	0.231	-0.0817
	(1.521)	(1.541)	(1.361)	(1.569)	(0.689)	(0.732)
Observations	35,039	33,061	38,511	37,761	26,889	23,864
R^2	0.829	0.846	0.793	0.807	0.710	0.715
Extended Cell by Year FE	N	Y	N	Y	N	Y
Dep. Var. Mean	477	477	380.3	380.3	99.35	99.35

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of an ever settling plant are included. All mortality rates are age-adjusted and expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder are excluded (<20 deaths per year). The CDC does not report death rates for counties below a given number of deaths—because there are fewer deaths for respiratory than cardiovascular death, there are fewer observations available for both respiratory deaths (and cardiovascular plus respiratory deaths) than cardiovascular deaths alone. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Counties in the same extended cells must also be both above/below median of the outcome variable at baseline. Standard errors are clustered at the state level.

Table C.10: Effect of Settlement on Cardiovascular, Respiratory and Other Mortality: Crude Mortality Rates

VARIABLES	(1) Ttl Card+Resp Mort Rate	(2) Ttl Card+Resp Mort Rate	(3) Ttl Card Mort Rate	(4) Ttl Card Mort Rate	(5) Ttl Resp Mort Rate	(6) Ttl Resp Mort Rate
Post Consent Decree w/i 40 Mi	-7.466***	-7.745***	-4.813**	-5.244**	-2.640***	-2.364***
	(2.206)	(2.014)	(2.113)	(2.041)	(0.811)	(0.687)
Post Consent Decree 40-80 Mi	-0.121	0.124	-0.242	0.188	0.251	0.119
	(2.008)	(1.987)	(1.973)	(1.965)	(0.937)	(0.902)
Observations	21,470	20,705	21,493	20,754	21,079	20,072
R^2	0.930	0.935	0.919	0.925	0.804	0.821
Extended Cell by Year FE	N	Y	N	Y	N	Y
Dep. Var Mean (1998)	467.8	468.8	364	367.6	110.5	110.5

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of a coal plant and with greater 25,000 population in 1998 are included. All mortality rates are expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder are excluded (<20 deaths in a year). All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Standard errors are clustered at the state level.

Table C.11: Effect of Settlement on Cardiovascular and Respiratory Mortality: Different Distance Bins

	(1) Card+Resp Mort Rate	(2) Card Mort Rate	(3) Resp Mort Rate	(4) Card+Resp Mort Rate	(5) Card Mort Rate	(6) Resp Mort Rate	(7) Card+Resp Mort Rate	(8) Card Mort Rate	(9) Resp
Post Consent Decree w/i 30 Mi	-4.469	-2.007	-2.341**						
	(2.762)	(2.088)	(1.047)						
Post Consent Decree w/i 30-60 Mi	-2.823	-2.517	-0.501						
	(1.911)	(1.767)	(0.672)						
Post Consent Decree w/i 50 Mi				-4.686**	-3.252*	-1.428			
				(2.280)	(1.767)	(1.007)			
Post Consent Decree w/i 50-100 Mi				-0.287	0.170	-0.306			
				(1.573)	(1.537)	(0.683)			
Post Consent Decree w/i 60 Mi							-4.414*	-2.974	-1.605*
							(2.222)	(1.936)	(0.810)
Post Consent Decree w/i 60-120 Mi							-0.779	0.351	-0.849
							(1.086)	(0.956)	(0.699)
Observations	19,540	19,544	19,202	22,616	22,631	22,192	23,389	23,412	22,939
R^2	0.897	0.890	0.698	0.893	0.887	0.685	0.892	0.886	0.680
Cell by Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 60 miles (cols (1)-(3)), 100 miles (Cols (4)-(6), or 120 miles (Cols (7)-(9)) of an ever settling plant are included. All mortality rates are age-adjusted and expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder are excluded (<20 deaths in a year). Only counties with greater than 25,000 population at baseline are included. Because cardiovascular deaths are more common than respiratory deaths, more counties have reliable cardiovascular death rates than respiratory death rates—this is why the cardiovascular mortality regressions have slightly more observations. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Standard errors are clustered at the state level.

Table C.12: Effect of Settlement on Cardiovascular and Respiratory Mortality: Single vs Multiple Plants Under Consent Decree

VARIABLES	(1) Ttl Card+Resp Mort Rate	(2) Ttl Card+Resp Mort Rate	(3) Ttl Card Mort Rate	(4) Ttl Card Mort Rate	(5) Ttl Resp Mort Rate	(6) Ttl Resp Mort Rate
1 Plant Post Consent Decree w/i 40 Mi	-5.628**	-5.154**	-3.525*	-3.245**	-2.110*	-1.960*
	(2.531)	(2.124)	(1.788)	(1.594)	(1.201)	(1.025)
>1 Plant Post Consent Decree w/i 40 Mile	-8.586**	-8.849**	-5.649*	-6.696*	-2.937***	-2.190**
	(3.691)	(3.748)	(3.061)	(3.705)	(1.062)	(0.927)
1 Plant Post Consent Decree w/i 40-80 Mi	0.505	0.923	0.741	1.330	-0.240	-0.163
	(1.890)	(2.190)	(1.642)	(2.175)	(0.952)	(0.833)
>1 Plant Post Consent Decree w/i 40-80 Mi	-0.272	0.418	-1.435	-0.512	1.159	1.051
	(2.675)	(2.987)	(2.189)	(2.456)	(1.094)	(1.034)
Observations	21,470	20,705	21,493	20,754	21,470	20,072
R^2	0.893	0.902	0.888	0.897	0.678	0.721
Extended Cell by Year FE	N	Y	N	Y	N	Y

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of a coal plant and with greater than 25,000 population in 1998 are included. All mortality rates are age-adjusted and expressed as rates per 100,000 population County-years with mortality rates deemed unreliable by CDC Wonder are excluded (<20 deaths). Because cardiovascular deaths are more common than respiratory deaths, more counties have reliable cardiovascular death rates than respiratory death rates—this is why the cardiovascular mortality regressions have slightly more observations. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Counties in the same extended cells must also be both above/below median of the outcome variable at baseline. Standard errors are clustered at the state level.

Table C.13: Effect of Settlement on Utility Price: Robustness Check

	(1) Ln Avg Price	(2) Avg Price	(3) Avg Price	(4) Avg Price	(5) Avg Price	(6) Avg Price	(7) Avg Price	(8) Avg Price	(9) Avg Price
Post Consent Decree	0.118***	0.00940**	0.00871***	0.0123**	0.00901**	0.00805***	0.00894**	0.0103***	0.00794**
	(0.0421)	(0.00369)	(0.00286)	(0.00486)	(0.00353)	(0.00282)	(0.00397)	(0.00332)	(0.00378)
Observations	1,727	1,666	1,685	1,656	1,606	1,330	2,429	2,750	13,632
R^2	0.870	0.885	0.889	0.831	0.836	0.905	0.955	0.869	0.938
Spec	Base	Base	Base	NERC x Year only	En Share Quart	Base	Base	Base	Base
Sample	Base	Balanced	6yr Pre/Post	Base	Cust>100K	Pre-2011	No Cust Cut-off	All Util Types,>25K Cust	All Util
Dep. Var Mean (1998)	0.0999	0.100	0.0999	0.0999	0.0951	0.100	0.128	0.100	0.113

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on utility average retail price measured in $/kWh. Only investor-owned utilities that both operate plants and distribute electricity and had greater than 25,000 population in 1998 are included unless otherwise noted. Settlement is considered the year of the utility's first New Source Review settlement. Coefficients come from a regression of the outcome variable on indicators for years relative to settlement (never-settling plants have a value of 0 for all years). All regressions include utility fixed effects and year by cell fixed effects unless otherwise noted. Cells consist of all utilities that are all above/below median 1998 customer level ,above/below median 1998 average price, above/below median 1998 coal share of generation and above/below median 1998 natural gas share of generation. Robustness specifications are as follows. In Column (1), the dependent variable is logged. In Column (2), only utilities with observations in all years are included. In Column (3), only 6 years before and after a settlement are included for treated utilities. In Column (4), controls include only NERC by Year and utility fixed-effects. In Column (5), cells include energy generation share quartiles instead of medians. In Column (6), only IOUs with >100,000 1998 customers are included. In Column (7), IOUs with <25,000 1998 customers are included. In Column (8), all publicly-owned utilities with >25,000 1998 customers are included. In Column (9), all utilities that generate and transmit electricty are included. In the latter three specifications, cells are interacted with a utility ownership type fixed effects. All standard errors are clustered at the utility level.

Table C.14: Effect of Settlement on Utility Revenue: Robustness Check

	(1) Ttl Rev	(2) Ln Ttl Rev	(3) Ln Ttl Rev	(4) Ln Ttl Rev	(5) Ln Ttl Rev	(6) Ln Ttl Rev	(7) Ln Ttl Rev	(8) Ln Ttl Rev	(9) Ln Ttl Rev
Post Consent Decree	187,327	0.121***	0.118***	0.179***	0.118***	0.101***	0.122*	0.130**	0.0858
	(122,382)	(0.0447)	(0.0358)	(0.0656)	(0.0434)	(0.0320)	(0.0692)	(0.0508)	(0.0593)
Observations	1,740	1,666	1,685	1,656	1,606	1,330	2,429	2,750	13,632
R^2	0.964	0.869	0.874	0.788	0.832	0.899	0.984	0.986	0.995
Spec	Base	Base	Base	NERC x Year only	En Share Quart	Base	Base	Base	Base
Sample	Base	Balanced	6yr Pre/Post	Base	Cust>100K	Pre-2011	No Cust Cut-off	All Util Types	All
Unlogged Dep. Var Mean (1998)	1.992e+06	2.084e+06	1.992e+06	1.992e+06	2.242e+06	1.966e+06	1.528e+06	1.329e+06	290925

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on log utility revenue. Only investor-owned utilities that both operate plants and distribute electricity and had greater than 25,000 population in 1998 are included unless otherwise noted. Settlement is considered the year of the utility's first New Source Review settlement. Coefficients come from a regression of the outcome variable on indicators for years relative to settlement (never-settling plants have a value of 0 for all years). All regressions include utility fixed effects and cell fixed effects. Cells consist of all utilities that are all above/below median 1998 customer level, above/below median 1998 average price, above/below median 1998 coal share of generation and above/below median 1998 natural gas share of generation unless otherwise noted. Robustness specifications are as follows. In Column (1), the dependent variable is in levels. In Column (2), only utilities with observations in all years are included. In Column (3), only 6 years before and after a settlement are included for treated utilities. In Column (4), controls include only NERC by Year and utility fixed-effects. In Column (5), cells include energy generation share quartiles instead of medians. In Column (6), only IOUs with >100,000 1998 customers are included. In Column (7), IOUs with <25,000 1998 customers are included. In Column (8), all publicly-owned utilities with >25,000 1998 customers are included. In Column (9), all utilities that generate and transmit electricty are included. In the latter three specifications, cells are interacted with a utility ownership type fixed effects. All standard errors are clustered at the utility level.

C.2 Alternative Empirical Specification Accounting for Potential Bias with Two-way fixed effects

Recent work by Goodman-Bacon (2018) and Callaway and Sant'Anna (2019) have shown that difference-in-differences models with time-varying treatment may be biased in certain circumstances. In this section, I use an alternative empirical specification for the main analyses in the paper to address these concerns. In these alternate specifications, I use the same cell-based approach as in the main specification. However, within each cell I identify the first year in which any unit is treated and call this the event year for the cell.[26] I then compare these first-treated units with units in the cell that are never treated (dropping any units that receive treatment after the first treated unit). In this way, every treated unit has only a pure never-treated counterfactual (the average values of all untreated units in its cell in each year), ensuring that earlier treated units are not serving as counterfactuals for later treated units. I show further robustness to using a balanced panel of openings with exactly six years of data before and after each cells' treated year. Both specifications have substantially less power than the main specifications used in the paper because I am abandoning one source of variation (settlement timing among units that ever settle). Nonetheless, the results remain broadly consistent with the paper's main findings.

Table C.15 show the unit-level pollution emission results under this alternative specification. Results are nearly identical to the main analysis with large reductions in NO_x and SO_2, which are driven by changes in the level of the emission rates. Table C.16 and C.17 show results for the air quality and mortality analyses. Again, results are very similar to those observed using our primary identification strategy. It is important to note that in these analyses I am using only treated units that have the earliest consent decree occurring within 40 miles of the monitor (county centroid) within their cell; I control for consent decrees occurring within 40-80 miles of the

[26] In the case of the air quality and mortality analyses, I use the first year that a unit had a consent decree within 40 miles of the monitor (county centroid). The 40-80 miles indicator remains the same as in the primary analysis in order to control for consent decrees outside the 40 mile radius, which could still potentially affect local air pollution/mortality rates.

monitor (centroid) in the same manner as in the primary analysis.. Finally, Table C.18 show the results for the incidence analysis under this new specification. As above, the results remain largely unchanged. Together these results provide additional reassurance that the main effects observed in the paper are truly caused by the plant's coming under consent decrees and not bias induced through the chosen empirical specification.

Table C.15: Effect of Settlement on Coal-Fired Power Plant: Pure Controls Only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NO_x Tons	NO_x Tons	SO_2 Tons	SO_2 Tons	NO_x Emis Rate	NO_x Emis Rate	SO_2 Emis Rate	SO_2 Emis Rate
	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Post Consent Decree x Treated Plant	-1,156**	-1,162**	-2,616**	-3,549*	-31.68***	-26.03***	-82.50	-103.3
	(450.4)	(540.1)	(1,140)	(2,036)	(12.06)	(8.059)	(53.81)	(98.34)
Observations	8,852	2,717	8,852	2,717	7,582	2,278	7,582	2,278
R^2	0.866	0.907	0.843	0.820	0.884	0.917	0.872	0.902
Sample	Full	Balanced	Full	Balanced	Full	Balanced	Full	Balanced
Dep. Var Mean (1998)	6084	6084	13524	13524	299.8	299.8	857.7	857.7

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of settlement on pollution and generation outcomes among coal-fired power plants. The number of observations for emission rates shrinks because I drop years with less than 1 MM BTU of heat input; rates become unstable with very low-levels of heat input. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input, below/above median SO_2 emission rates, and above/below median NO_x emission rate Only treated units whose settlement years were the first in their cell and never-treated units are included in order to create a pure control group. A "balanced" sample includes only monitors whose cell's settlement year was between 2004 and 2007, which allows for 6 years of pre and 6 years of post data. Pollution and generation data come from CEMS and nonattainment data come from the EPA's Green Book. Regressions are weighted by a unit's share of a plant's heat input in 1998, the baseline year. Standard errors are clustered at the state level.

Table C.16: Effect of Settlement on Ambient Pollution Values: Pure Controls Only

	(1)	(2)	(3)	(4)	(5)	(6)
	SO_2 ppb	SO_2 ppb	Ozone ppm	Ozone ppm	PM 2.5 $\mu g/m^3$	PM 2.5 $\mu g/m^3$
Post Consent Decree x Treated Unit	-5.216***	-6.720***	0.000539	0.000901	-0.132	-0.224
	(1.770)	(1.506)	(0.000444)	(0.00173)	(0.180)	(0.232)
Observations	1,703	378	5,796	491	4,276	533
R^2	0.948	0.958	0.929	0.911	0.971	0.945
Sample	Full	Balanced	Full	Balanced	Full	Balanced
Dep. Var Mean (1998)	0.0900	0.0900	2.548e+06	2.548e+06	1.045e+06	1.045e+06

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on observed pollution levels in monitors nearby the plant. Only monitors that began operation prior to 1999, were still in operation in 2017 were included and had a plant under consent decree within 80 miles from the monitor were included. Ozone is measured in parts per million and is the annual average maximum 8 hour value. Sulfur dioxide is measured in parts per billion and is the average annual hourly value. Coefficients come from a regression of the outcome variable on indicators for whether or not a monitor has a plant under consent decree within 40 miles and and within 40-80 miles. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Only treated monitors whose settlement years were the first in their cell and never-treated monitors are included in order to create a pure control group. A "balanced" sample includes only monitors whose cell's settlement year was between 2004 and 2007, which allows for 6 years of pre and 6 years of post data. Air monitor data come from the EPA's AQS datamart. Standard errors are clustered at the monitor level.

Table C.17: Effect of Settlement on Cardiovascular, Respiratory and Other Mortality: Pure Controls Only

VARIABLES	(1) Ttl Card+Resp Mort Rate	(2) Ttl Card+Resp Mort Rate	(3) Ttl Card Mort Rate	(4) Ttl Card Mort Rate	(5) Ttl Resp Mort Rate	(6) Ttl Resp Mort Rate
Post Consent Decree w/i 40 Mi	-5.917**	-4.128	-4.125*	-2.084	-2.846**	-2.003
	(2.713)	(2.958)	(2.205)	(2.321)	(1.190)	(1.306)
Post Consent Decree w/i 40-80 Mi	0.812	1.906	-0.115	-2.762	1.134	0.990
	(2.516)	(3.026)	(2.742)	(2.211)	(1.095)	(1.470)
Observations	17,849	5,748	18,272	6,826	17,613	5,928
R^2	0.908	0.921	0.902	0.909	0.719	0.759
Sample	Full	Balanced	Full	Balanced	Full	Balanced

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on mortality rates in counties near plants with and without consent decrees. Only counties whose centroids are within 80 miles of an ever settling plant are included. County-years with mortality rates deemed unreliable by CDC Wonder are excluded. All mortality rates are per 100,000 population and are age-adjusted. The CDC does not report death rates for counties below a given number of deaths—because there are fewer deaths for respiratory than cardiovascular death, there are fewer observations available for both respiratory deaths (and cardiovascular plus respiratory deaths) than cardiovascular deaths alone. All regressions include unit fixed effects, cell by year fixed effects, controls for NAAQS non-attainment status and NO_x Budget Trading Program by year fixed effects. Cells are defined as all units in the same state, below/above median 1998 heat input within 40 miles and below/above median 1998 heat input within 40-80 miles. Counties in the same extended cells must also be both above/below median of the outcome variable at baseline. Mortality rate data comes from the CDC Wonder database. Standard errors are clustered at the state level.

Table C.18: Effect of Settlement on Ln Average Electricity Prices, Ln Utility Revenue and Ln Total Customers: Pure Controls Only

	(1) Ln Avg Price	(2) Ln Avg Price	(3) Ln Tot Rev	(4) Ln Tot Rev	(5) Ln Tot Cust	(6) Ln Tot Cust
Post Consent Decree x Treated Unit	0.167***	0.0931*	0.177**	0.119	0.0244	
	(0.0506)	(0.0517)	(0.0678)	(0.0746)	(0.0378)	
Observations	1,003	299	1,003	299	610	299
R^2	0.810	0.772	0.961	0.979	0.987	0.998
Sample	Full	Balanced	Full	Balanced	Full	Balanced
Unlogged Dep. Var Mean (1998)	0.0900	0.0900	2.548e+06	2.548e+06	1.045e+06	1.045e+06

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows the effect of a settlement on log average electricity price, log utility revenue and log total customers. Only investor-owned utilities that both operate plants and distribute electricity and had greater than 25,000 population in 1998 are included unless otherwise noted. Settlement is considered the year of the utility's first New Source Review settlement. Coefficients come from a regression of the outcome variable on indicators for years relative to settlement (never-settling plants have a value of 0 for all years). All regressions include utility fixed effects and cell fixed effects. Cells consist of all utilities that are are both above/below median 1998 customer level ,above/below median 1998 average price, above/below median 1998 avg annual energy usage, above/below median 1998 coal share of generation and above/below median 1998 natural gas share of generation. Utility data come from EIA Form 861. All standard errors are clustered at the utility level.

Chapter 4

Bibliography

Bibliography

[1] Anna Aizer and Joseph J Doyle Jr. "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges". In: *The Quarterly Journal of Economics* 130.2 (2015), pp. 759–803.

[2] Nathan B Anderson. "Beggar thy neighbor? Property taxation of vacation homes". In: *National Tax Journal* (2006), pp. 757–780.

[3] Spencer Banzhaf, Lala Ma, and Christopher Timmins. "Environmental justice: The economics of race, place, and pollution". In: *Journal of Economic Perspectives* 33.1 (2019), pp. 185–208.

[4] Alan I Barreca, Matthew Neidell, and Nicholas J Sanders. "Long-run pollution exposure and adult mortality: Evidence from the acid rain program". In: (2017).

[5] Geoffrey Barrows, Teevrat Garg, and Akshaya Jha. "The economic benefits versus environmental costs of India's coal fired power plants". In: *Available at SSRN 3281904* (2018).

[6] Alexander Bartik, Janet Currie, Michael Greenstone, and Christopher R Knittel. "The local economic and welfare consequences of hydraulic fracturing". In: *Available at SSRN 2692197* (2018).

[7] Patrick Bayer, Fernando Ferreira, and Robert McMillan. "A unified framework for measuring preferences for schools and neighborhoods". In: *Journal of Political Economy* 115.4 (2007), pp. 588–638.

[8] Gary S Becker. "Crime and punishment: An economic approach". In: *The Economic Dimensions of Crime*. Springer, 1968, pp. 13–68.

[9] Manudeep Bhuller, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. "Incarceration, recidivism and employment". In: (2016).

[10] Barbara Biasi. "School finance equalization increases intergenerational mobility: Evidence from a simulated-instruments approach". In: (2019).

[11] Sandra E Black. "Do better schools matter? Parental valuation of elementary education". In: *The Quarterly Journal of Economics* 114.2 (1999), pp. 577–599.

[12] Wesley Blundell, Gautam Gowrisankaran, and Ashley Langer. “Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations”. In: *American Economic Review* (Forthcoming).

[13] J Tom Boer, Manuel Pastor, James L Sadd, and Lori D Snyder. “Is there environmental racism? The demographics of hazardous waste in Los Angeles County”. In: *Social Science Quarterly* 78.4 (1997), pp. 793–810.

[14] Judson Boomhower. “Drilling like there’s no tomorrow: Bankruptcy, insurance, and environmental risk”. In: *American Economic Review* 109.2 (2019), pp. 391–426.

[15] Severin Borenstein and James B Bushnell. “Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency”. In: (2018).

[16] Daria Burnes, David Neumark, and Michelle J White. “Fiscal zoning and sales taxes: do higher sales taxes lead to more retailing and less manufacturing?” In: (2011).

[17] James B Bushnell and Catherine D Wolfram. “Enforcement of vintage differentiated regulations: The case of new source review”. In: *Journal of Environmental Economics and Management* 64.2 (2012), pp. 137–152.

[18] Brantly Callaway and Pedro HC Sant’Anna. “Difference-in-differences with multiple time periods”. In: *Available at SSRN 3148250* (2019).

[19] Sebastian Calonico, Matias D Cattaneo, Max H Farrell, and Rocıo Titiunik. “rdrobust: Software for regression discontinuity designs”. In: *Stata Journal* 17.2 (2017), pp. 372–404.

[20] David Card and A Abigail Payne. “School finance reform, the distribution of school spending, and the distribution of student test scores”. In: *Journal of Public Economics* 83.1 (2002), pp. 49–82.

[21] Maite Careaga and Barry Weingast. “Fiscal federalism, good governance, and economic growth in Mexico”. In: *In Search of Prosperity: Analytical Narratives on Economic Growth* (2003), pp. 399–435.

[22] Elizabeth U Cascio and Ebonya Washington. “Valuing the vote: The redistribution of voting rights and state funds following the voting rights act of 1965”. In: *The Quarterly Journal of Economics* 129.1 (2013), pp. 379–433.

[23] Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. “The value of school facility investments: Evidence from a dynamic regression discontinuity design”. In: *The Quarterly Journal of Economics* 125.1 (2010), pp. 215–261.

[24] Aaron Chalfin and Justin McCrary. “Criminal deterrence: A review of the literature”. In: *Journal of Economic Literature* 55.1 (2017), pp. 5–48.

[25] H Ron Chan and Yichen Zhou. “Regulatory spillover and climate co-benefits: Evidence from the New Source Review lawsuits”. In: *Available at SSRN 3399713* (2019).

[26] CJ Ciaramella. “Virginia suspends hundreds of thousands of driver’s licenses due to unpaid court debt”. In: *Reason* (2016).

[27] Steve Cicala. “When does regulation distort costs? lessons from fuel procurement in us electricity generation”. In: *American Economic Review* 105.1 (2015), pp. 411–44.

[28] RR Cirillo, TD Wolsko, RO Mueller, PA Dauzvardis, MJ Senew, K Gamauf, and DA Seymour. “Evaluation of regional trends in power plant siting and energy transport”. In: (1977).

[29] *City of Redmond v. Moore*. 2004.

[30] Mark Clayton. “EPA’s record settlement with utility could lead to other deals”. In: *Christian Science Monitor* (2007). URL: https://www.csmonitor.com/2007/1010/p25s04-usgn.html.

[31] Janet Currie, Lucas Davis, Michael Greenstone, and Reed Walker. “Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings”. In: *American Economic Review* 105.2 (2015), pp. 678–709.

[32] Lucas W Davis. “The effect of power plants on local housing values and rents”. In: *Review of Economics and Statistics* 93.4 (2011), pp. 1391–1402.

[33] Matthew Davis and Fernando V Ferreira. “Housing disease and public school finances”. In: (2017).

[34] Arindrajit Dube, T William Lester, and Michael Reich. “Minimum wage effects across state borders: Estimates using contiguous counties”. In: *The Review of Economics and Statistics* 92.4 (2010), pp. 945–964.

[35] *Environmental Defense v. Duke Energy Corporation*. 2007.

[36] Steven Ferrey. “Siting technology, land-Use energized”. In: *Catholic University Law Review* 66 (2016), p. 1.

[37] William A Fischel. “The congruence of American school districts with other local government boundaries: A Google-Earth exploration”. In: *Available at SSRN 967399* (2010).

[38] Meredith Fowlie. “Emissions trading, electricity restructuring, and investment in pollution abatement”. In: *American Economic Review* 100.3 (2010), pp. 837–69.

[39] Lucie Gadenne and Monica Singhal. “Decentralization in developing economies”. In: *Annual Review of Economics* 6.1 (2014), pp. 581–604.

[40] GAO. “EPA needs better information on New Source Review permits”. In: (2012).

[41] Markus Gehrsitz. “Speeding, punishment, and recidivism: Evidence from a regression discontinuity design”. In: *The Journal of Law and Economics* 60.3 (2017), pp. 497–528.

[42] Andrew Gelman and Guido Imbens. “Why high-order polynomials should not be used in regression discontinuity designs”. In: *Journal of Business & Economic Statistics* (2018), pp. 1–10.

[43] Stephen Gibbons. “Gone with the wind: Valuing the visual impacts of wind turbines through house prices”. In: *Journal of Environmental Economics and Management* 72 (2015), pp. 177–196.

[44] Felipe Goncalves and Steven Mello. “Does the punishment fit the crime? Speeding fines and recidivism”. In: (2017).

[45] Andrew Goodman-Bacon. “Difference-in-differences with variation in treatment timing”. In: (2018).

[46] Wayne B Gray and Jay P Shimshack. “The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence”. In: *Review of Environmental Economics and Policy* 5.1 (2011), pp. 3–24.

[47] Michael Greenstone, Richard Hornbeck, and Enrico Moretti. “Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings”. In: *Journal of Political Economy* 118.3 (2010), pp. 536–598.

[48] Joseph Gyourko, Albert Saiz, and Anita Summers. “A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index”. In: *Urban Studies* 45.3 (2008), pp. 693–729.

[49] Li Han and James Kai-Sing Kung. “Fiscal incentives and policy choices of local governments: Evidence from China”. In: *Journal of Development Economics* 116 (2015), pp. 89–104.

[50] Benjamin Hansen. “Punishment and deterrence: Evidence from drunk driving”. In: *American Economic Review* 105.4 (2015), pp. 1581–1617.

[51] Garth Heutel. “Plant vintages, grandfathering, and environmental policy”. In: *Journal of Environmental Economics and Management* 61.1 (2011), pp. 36–51.

[52] Stephen P Holland, Erin T Mansur, Nicholas Muller, and Andrew J Yates. “Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation”. In: (2018).

[53] Caroline M Hoxby. “All school finance equalizations are not created equal”. In: *The Quarterly Journal of Economics* 116.4 (2001), pp. 1189–1231.

[54] Caroline M Hoxby and Ilyana Kuziemko. “Robin Hood and his not-so-merry plan: Capitalization and the self-destruction of Texas’ school finance equalization plan”. In: (2004).

[55] C Kirabo Jackson, Rucker Johnson, and Claudia Persico. “The effect of school finance reforms on the distribution of spending, academic achievement, and adult outcomes”. In: (2014).

[56] C Kirabo Jackson, Rucker C Johnson, and Claudia Persico. “The effects of school spending on educational and economic outcomes: Evidence from school finance reforms”. In: *The Quarterly Journal of Economics* 131.1 (2015), pp. 157–218.

[57] Reid Johnson, Jacob LaRiviere, and Hendrik Wolff. “Fracking, coal, and air quality”. In: (2017).

[58] Elena Kantorowicz-Reznichenko. “Day-Fines: Should the rich pay more?” In: *Review of Law & Economics* 11.3 (2015), pp. 481–501.

[59] Nathaniel O Keohane, Erin T Mansur, and Andrey Voynov. “Averting regulatory enforcement: Evidence from new source review”. In: *Journal of Economics & Management Strategy* 18.1 (2009), pp. 75–104.

[60] Julien Lafortune, Jesse Rothstein, and Diane Whitmore Schanzenbach. “School finance reform and the distribution of student achievement”. In: *American Economic Journal: Applied Economics* 10.2 (2018), pp. 1–26.

[61] Julien Lafortune and David Schonholzer. “School finance reform and the distribution of student achievement”. In: *Working Paper* (2019).

[62] Ian Lange and Joshua Linn. “Bush v. Gore and the effect of new source review on power plant emissions”. In: *Environmental and Resource Economics* 40.4 (2008), pp. 571–591.

[63] David S Lee and Thomas Lemieux. “Regression discontinuity designs in economics”. In: *Journal of Economic Literature* 48.2 (2010), pp. 281–355.

[64] The Legal Aid Justice Center. “Driving on empty: Payment plan reforms don’t fix Virginia’s court debt crisis”. In: *Technical Report* (2018).

[65] Steven D Levitt. “Incentive compatibility constraints as an explanation for the use of prison sentences instead of fines”. In: *International Review of Law and Economics* 17.2 (1997), pp. 179–192.

[66] John A List, Daniel L Millimet, Per G Fredriksson, and W Warren McHone. “Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator”. In: *Review of Economics and Statistics* 85.4 (2003), pp. 944–952.

[67] Simon Luechinger. “Air pollution and infant mortality: A natural experiment from power plant desulfurization”. In: *Journal of Health Economics* 37 (2014), pp. 219–231.

[68] Byron Lutz. “Quasi-experimental evidence on the connection between property taxes and residential capital investment”. In: *American Economic Journal: Economic Policy* 7.1 (2015), pp. 300–330.

[69] Michael D Makowsky, Thomas Stratmann, and Alexander T Tabarrok. “To serve and collect: The fiscal and racial determinants of law enforcement”. In: (2018).

[70] Joseph Marchand, Jeremy Weber, et al. "The labor market and school finance effects of the Texas shale boom on teacher quality and student achievement". In: (2015).

[71] Luis R Martinez. "Sources of revenue and government performance: evidence from Colombia". In: *Available at SSRN 3273001* (2016).

[72] Thomas O McGarity. "When Strong enforcement works better than weak regulation: The EPA/DOJ New Source Review Enforcement Initiative". In: *Maryland Law Review* 72 (2012), p. 1204.

[73] Steve Mello. "Speed traps or poverty traps: The effects of nuisance policing on financial health". In: *Princeton University Job Market Paper* (2018).

[74] Amalia R Miller and Carmit Segal. "Does temporary affirmative action produce persistent effects? A study of black and female employment in law enforcement". In: *Review of Economics and Statistics* 94.4 (2012), pp. 1107–1125.

[75] Corbin L Miller. "The Effect of education spending on student achievement: Evidence from property values and school finance rules". In: (2018).

[76] Michael Mueller-Smith. "The criminal and labor market impacts of incarceration". In: *Working Paper* (2015).

[77] Michael Mueller-Smith and Kevin T Schnepel. "Punishment and (non-)deterrence: Evidence on first-time drug offenders from regression discontinuities". In: (2016).

[78] Daniel S Nagin. "Deterrence: A review of the evidence by a criminologist for economists". In: *Annual Review of Economics* 5.1 (2013), pp. 83–105.

[79] Jonathan Remy Nash and Richard L Revesz. "Grandfathering and environmental regulation: the law and economics of new source review". In: *Northwestern University Law Review* 101 (2007), p. 1677.

[80] Phuong Nguyen-Hoang and John Yinger. "The capitalization of school quality into house values: A review". In: *Journal of Housing Economics* 20.1 (2011), pp. 30–48.

[81] Wallace E Oates. "The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis". In: *Journal of Political Economy* 77.6 (1969), pp. 957–971.

[82] Wallace E Oates and William A Fischel. "Are local property taxes regressive, progressive, or what?" In: *National Tax Journal* 69.2 (2016), p. 415.

[83] Claudia Persico and Joanna Venator. "The effects of local industrial pollution on students and schools". In: (2018).

[84] A Mitchell Polinsky. "The optimal use of fines and imprisonment when wealth is unobservable". In: *Journal of Public Economics* 90.4-5 (2006), pp. 823–835.

[85] Justin M Ross. “Are community-nuisance fiscal zoning arrangements undermined by state property tax reforms? Evidence from nuclear power plants and school finance equalization”. In: *Land Economics* 89.3 (2013), pp. 449–465.

[86] Tracy Samilton. “Monroe County says DTE Energy blindsided it in property tax negotiation”. In: (2018).

[87] Michael W Sances and Hye Young You. “Economic shocks, targeted transfers, and local public goods: Evidence from US shale gas boom”. In: (2017).

[88] Joseph S Shapiro and Reed Walker. “Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade”. In: *American Economic Review* 108.12 (2018), pp. 3814–54.

[89] Megan T Stevenson and Sandra G Mayson. “The scale of misdemeanor justice”. In: (2018).

[90] M Weber, A Srikanth, and B Baker. “School funding fairness data system codebook”. In: *Newark, NJ: Rutgers Graduate School of Education-Education Law Center* (2016).

[91] Barry R Weingast. “Second generation fiscal federalism: The implications of fiscal incentives”. In: *Journal of Urban Economics* 65.3 (2009), pp. 279–293.

[92] Nat Williams. “Rural school dependent on power plant”. In: (2018).

[93] *Wisconsin Elec. Power Co. v. Reilly*. 1990.

[94] Ekaterina V Zhuravskaya. “Incentives to provide local public goods: fiscal federalism, Russian style”. In: *Journal of Public Economics* 76.3 (2000), pp. 337–368.