ESSAYS ON INTERGOVERNMENTAL GRANT PROGRAMS

By

Patrick J. Flynn

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in

Economics

May 13, 2022

Nashville, Tennessee

Approved:

Lesley Turner, Ph.D.

Michelle Marcus, Ph.D.

Andrew Goodman-Bacon, Ph.D.

Christopher Carpenter, Ph.D.

To my father, Joe, my mother, Mary Lou, and my friends, Kay, Q, Dave, Sarah, CJ, Bacon and Sayeh.

ACKNOWLEDGMENTS

For their guidance, I thank my dissertation committee, Lesley Turner, Michelle Marcus, Andrew Goodman-Bacon and Kitt Carpenter. I also thank Anna Aizer, Matt Baird, Brian Beach, Lindsey Bullinger, Analisa Packham, Sayeh Nikpay, Matthew Pesner, Pedro Sant'Anna, Tucker Smith, David Wehrly, Will Wheeler and Barton Willage for their valuable insights and their support of my research. I am grateful for feedback from seminar participants at the Southern Economic Association Annual Conference, the Missouri Valley Economic Association Annual Conference, and the Association of Environmental and Resource Economists Summer Conference. Research in this dissertation was enhanced via financial support from the Kirk Dornbush Summer Research Grant. This dissertation would not have been possible if not for support and encouragement from Kayleigh McCrary, Colin Sharpe, Zizheng Luo, Ji Hye Heo, Lachlan Watkins and my other classmates, as well as the faculty in the Vanderbilt Economics Department and the staff at the Tennessee Education Research Alliance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Rivers, Lakes and Revenue Streams: The Heterogeneous Effects of Clean Water Act Grants on Local Spending

From 1972 to 1988, the Clean Water Act funded $153 billion (in 2014 dollars) in grants to municipal governments for capital upgrades to wastewater treatment facilities. While CWA grants caused significant reductions in pollution and increases in housing prices (Keiser and Shapiro, 2019a), as well as improvements in infant health (Flynn and Marcus, 2021), analyses of the CWA consistently estimate benefit/cost ratios below one (Keiser et al., 2019) with each grant dollar generating an average return of $0.45 (Keiser and Shapiro, 2019a; Flynn and Marcus, 2021). Intergovernmental transfers like CWA grants account for around one-fifth of all municipal government revenues (Urban-Brookings Tax Policy Center, 2017), but donor governments often have little control over how receiving governments adjust spending in response to grant receipt. In most cases, having a high rate of "grant pass-through", the increase in targeted expenditure induced by a dollar of grant revenue, is a necessary condition for grant programs to accomplish their policy goals.

This paper uses recently developed difference-in-differences methods (Deshpande and Li, 2019; Callaway and Sant'Anna, 2020b; Callaway et al., 2021) to show that, given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded wastewater capital upgrades. Our analysis combines data on every CWA grant with spending data from 112 municipal governments. We compare changes in spending in newly treated municipalities to changes in spending in municipalities that will be treated in future periods to show that, on average, each dollar of grant revenue is associated with a $0.45 increase in sewerage capital spending. Dividing previously estimated benefit to cost ratios of CWA grants from Keiser and Shapiro (2019a) and Flynn and Marcus (2021) by this pass-through estimate suggests that each CWA grant dollar that municipalities spent on sewerage capital generated an average return of $1.01.

Along with funding grants, the Act imposed new capital standards on all wastewater treatment facilities in the United States. Municipalities could use grant funds to offset the new costs that the CWA's capital mandate imposed on them, but many CWA grants were substantially larger than the cost of coming into compliance with the mandate. This motivates an examination of how pass-through relates to mandated costs. Using data from an EPA engineering assessment of wastewater capital upgrade costs, we divide our sample of municipalities into those that received grants that were approximately equal to mandated costs, and those that received grants that were larger than mandated costs. We show that grants caused a dollar-for-dollar increase in sewerage capital spending in municipalities that received grants that were approximately equal to mandated costs, but in municipalities that received grants that were larger than mandated costs, sewerage capital spending only increased by 27 cents for every dollar of grant funding. The EPA also distributed many CWA grants to municipalities with wastewater treatment facilities that were already in compliance with the CWA's capital mandate. The mandate imposed no costs on these "compliant" municipalities, and they only increased sewerage capital spending by 3 cents for every dollar of grant funding.

The heterogeneous effects of CWA grants on local spending inform the design of intergovernmental grant programs and our understanding of the Clean Water Act. By providing evidence of full pass-through in municipalities where the CWA's capital mandate was binding, our results suggest that pairing grants with new regulation can increase the likelihood of full pass-through. This result may be generalizable to other intergovernmental transfer programs, such as grants for capital expenditure on highways and other infrastructure. This method of enacting policy is limited by the ability of donor governments to identify which receiving governments need fiscal assistance and estimate the costs of grant-funded programs. Both of these challenges to implementation hindered the CWA, leading to low grant pass-through.

Heterogeneity in pass-through across receiving municipalities can explain why our estimates are lower than previous estimates of CWA grant pass-through. In particular, Keiser and Shapiro (2019a) leverages continuous variation in grant size in a difference-in-differences framework to show evidence of nearly full pass-through of CWA grants. This type of estimate is identified in part off of comparisons of municipalities that receive large grants to municipalities that receive small grants before and after grant receipt, but in order to identify a meaningful result using this type of comparison, we must assume that the effect of a grant of a given size is homogeneous across units (Callaway et al., 2021). Because compliant and non-compliant municipalities respond to grant receipt differently, this assumption does not hold, and observations of compliant municipalities after grant receipt will not make valid counterfactuals for non-compliant municipalities (and vice-versa). We show that this causes estimates of CWA grant pass-through identified off of variation in grant size that do not account for heterogeneity in pass-through to over-estimate the effect of grant funding on targeted spending.

Similar problems can arise whenever there is heterogeneity in the relationship between grant amount and targeted spending across receiving municipalities. Evaluations of intergovernmental grant programs often estimate pass-through using difference-in-differences designs that leverage continuous variation in grant size, but pass-through is likely heterogeneous for many grant programs. For example, grant size is often positively correlated with receiving governments' preferences for public goods (Knight, 2002), so in many cases, the effect of a dollar of grant funding on targeted spending is increasing in grant size. This selection into grant size will bias pass-through estimates identified off of comparisons of units that receive large grants to units that receive small grants.

Finally, low pass-through in compliant municipalities motivates an examination of how these municipalities adjusted spending in response to grant receipt. CWA grants differ from many other intergovernmental transfers in that they support a function of local government that has a direct "fee-for-service" revenue stream where water and sewerage utilities collect charges from local residents. We show that compliant municipalities redistributed grant revenue to residents by reducing water and sewerage charges. This suggests that grant pass-through may be lower in settings where receiving governments control a channel through which they can redistribute money with relatively little friction and where costs of the public good are more salient to local residents than their benefits. Municipalities receiving federal grants for electrical utilities (USDA, 2019) or waste management (USDA, 2020), as well as those receiving modern wastewater treatment grants (Travis et al., 2004) may be able to use the same crowd-out mechanism.

Since annual historical spending data only exists for municipalities with populations over 75,000, we estimate pass-through for a relatively small number of municipalities. The municipalities in our sample are large, so our results represent the pass-through of $25.5 billion (2014$) of $153 billion total CWA grants. While we cannot directly estimate pass-through of CWA grants distributed to smaller municipalities, we provide evidence that municipal population is not correlated with pass-through for municipalities in our sample, which suggests that CWA grants outside of our sample may have a similar pass-through rate to the grants in our sample.

1.1 Background

1.1.1 The Clean Water Act

Congress originally passed The Clean Water Act in 1948 as the Federal Water Pollution Control Act. Facing pressure to enact policy to reduce surface water pollution after a series of high profile river fires, Congress significantly expanded the CWA in 1972. The strengthened CWA utilized both subsidies and new regulation to combat water pollution, distributing an estimated $153 billion (in 2014 dollars) in grants to fund upgrades to municipal wastewater treatment capital and imposing new capital standards on all municipal wastewater treatment plants in the United States.

The CWA funded wastewater treatment grants as follows. First, the EPA distributed grant money to states according to a congressionally mandated formula that was based on each state's wastewater treatment needs, as well as its total population and forecast population (Rubin, 1985). States then issued grants to municipalities according to priority lists that were based on the severity of nearby surface water pollution, the size of the population affected, the need for conservation of the affected waterway and that waterway's specific category of need.[1] States wrote their own priority lists, but these lists had to be approved by the EPA

[1] We show that this did not lead to selection on treated potential outcomes in Appendix Section 1.6.1.1

annually. The Clean Water Act also explicitly prohibited states from considering the geographic location of a receiving municipality, future population growth projections and any of a municipality's development needs that were not directly related to pollution abatement when writing priority lists (USEPA, 1980). Notably, federal and state governments did not explicitly allocate CWA grants according to the finances or spending preferences of receiving governments.

Many CWA grants were specifically designated for upgrades to wastewater treatment technology. When the CWA came into effect, more than a quarter of all wastewater treatment facilities in the US were using relatively inexpensive primary treatment (USEPA, 2000), which filters out large detritus and improves the aesthetics of surface water, but discharges all but the heaviest organic material into rivers and lakes (USEPA, 1998). The CWA required all wastewater treatment facilities to upgrade to more effective secondary treatment technology, which removes about 85 percent of organic material before discharge, by 1977.[2] Additionally, many states required facilities to satisfy treatment technology requirements that were more stringent than the CWA's mandate (USEPA, 2000).[3]

While the benefits of upgrading a facility's treatment technology were well understood, large upfront capital costs often made these upgrades prohibitively expensive without federal support. Upgrading could cost as much as 30 percent of the initial cost of the facility (National Environmental Research Center, 1972), so in general, municipalities only invested in secondary treatment before the CWA came into effect because of either state level regulation or pressure from nearby communities to reduce the flow of harmful pollutants downstream (Jerch, 2018).

Many CWA grants were larger than the cost of upgrading treatment technology.[4] In an assessment of the wastewater treatment upgrades mandated by the CWA, the EPA estimated the cost of upgrading to secondary treatment as a function of the amount of wastewater flowing through a facility (USEPA, 1973). We use these estimates to calculate the cost of upgrading for each non-compliant facility that we have wastewater flow data. Figure 1.1 plots the total costs that each municipality faced against its total grant amount. This figure shows that, while grants were similar to costs for municipalities that received up to around $125 per capita in total grant aid, grants to municipalities that received more than this amount were substantially larger than the

[2] The EPA enforced the secondary treatment standard through the National Pollutant Discharge Elimination System (NPDES), the CWA's new permitting system, which required more than 65,000 industrial and municipal dischargers to obtain permits from the EPA or state governments. Permits required municipal treatment plants to employ secondary treatment and had to be renewed every five years. Violating the terms of a permit resulted in a compliance order or civil suit by the EPA, and violators could be fined up to $25,000 per day (Copeland, 2016).

[3] Our sample includes some facilities that were in compliance with the CWA's capital mandate but were not in compliance with more stringent state standards. These facilities still faced incentives to spend CWA grants on upgrades, so we classify them as non-compliant. In Appendix Section 1.6.4.2, we show that our results are robust to dropping municipalities bound by state level regulations that were more stringent than the CWA's.

[4] States distributed grant dollars above the cost of upgrading to fund non-mandated capital expenditures, such as expansions of facility capacity.

estimated cost of upgrading.[5]

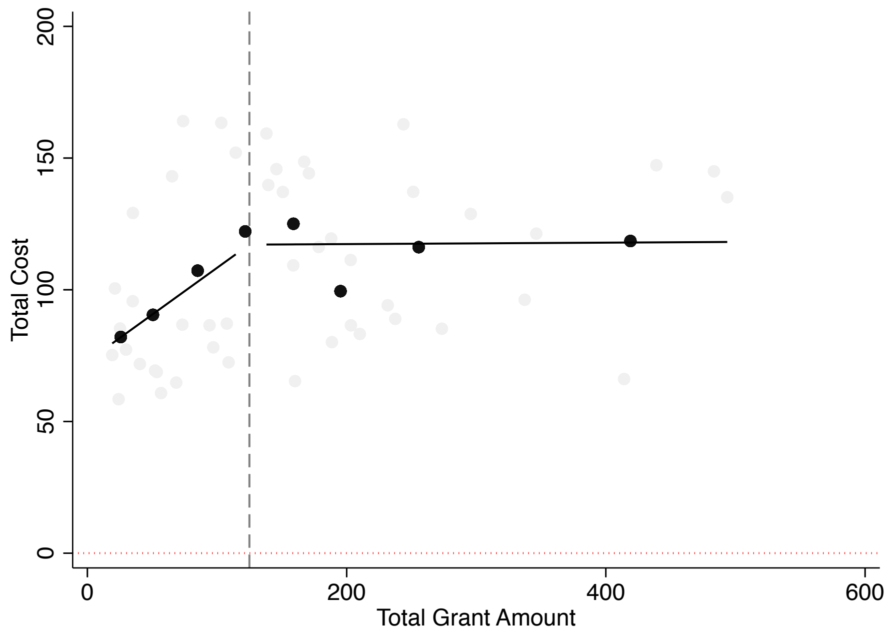


Figure 1.1: Costs Are Increasing in Grant Size Up to $125 and Constant Above $125

Notes: This figure plots estimates of non-compliant municipalities' total costs of upgrading to secondary treatment technology against the total grant dollars distributed to each municipality. The light gray dots are the true data and the black dots represent bins defined by deciles of total grant amount. Our cost estimates under-estimate costs for small grants, so we drop municipalities in the bottom decile of total grant aid. We divide the data at $125 per capita of total grant dollars, which is the point where costs are no longer increasing in grant size that we estimate in Table 1.3, and fit lines to the true data in each group.

Twelve percent of CWA grant dollars went to facilities that were already in compliance with both state and federal treatment technology standards. These grants were intended to fund non-mandated capital expenditures at wastewater facilities, such as capacity expansions.[6]

To illustrate variation in compliance and grant purpose, consider the example of two cities along the Willamette River in Oregon, historically one of the most polluted rivers in the Pacific Northwest. Eugene, OR is located at the top of the Willamette, while Portland lies at its base. To address municipal and industrial waste flowing downstream to Portland, the Oregon State Sanitary Authority (est. 1939) required Eugene and other municipalities discharging waste into the river to upgrade to secondary treatment in 1960. Because Eugene was bound by this state-level regulation, the city already met CWA treatment technology standards when they went into effect in 1972. Since Portland is located at the base of the Willamette, the city's discharge did not pollute other localities further downstream. Consequently, Portland was not subject to the 1960

[5]Figures 1.8 and 1.9 apply quadratic and linear fits to this data.

[6]The CWA allowed for grants to reimburse local governments that upgraded treatment technology prior to 1972 at a maximum 55% federal matching rate (USGAO, 1994), however, our grant data indicates that none of the complaint municipalities in our sample received grants for the purpose of reimbursement.

treatment technology requirement and did not satisfy CWA treatment technology standards in 1972. Portland received CWA grant funding to upgrade its treatment facility to meet new federal standards in 1974. Although their facility already satisfied the CWA's capital mandate, Eugene received federal funding to expand facility capacity and upgrade one of its pump stations.

1.1.2 Grant Pass-Through

Despite its importance to policy design, researchers and policymakers alike remain unsure of when to expect grant programs to have higher or lower pass-through rates.[7] Economic theory predicts that local governments should treat inframarginal grant revenue (i.e. grant revenue smaller than existing local expenditure) as if it is an unconditional lump sum transfer (Oates, 1999), but empirical results suggest that this is not often the case, and that, even in the absence of binding constraints, inframarginal transfer revenue "sticks where it hits" (Henderson, 1968; Gramlich, 1969).[8] Known as the "flypaper effect," this phenomenon appears in a range of settings.[9] Proposed explanations for the flypaper effect include mis-perception by voters (Courant et al., 1979), mis-specification by researchers (Megdal, 1987), uncertainty (Vegh and Vuletin, 2015) and a focus by researchers on the short- rather than long-term effects of revenue changes on local spending (Helm and Stuhler, 2021), but a consensus has yet to emerge (Inman, 2008).

CWA grants did not place any explicit restrictions on local spending. For this reason, although CWA grants were allocated for specific projects, a municipality could redistribute these conditional grant funds toward other purposes if they were inframarginal to that municipality's budget. The median baseline sewerage expenditure in our sample (\$20 per capita) was larger than the median yearly grant (\$15 per capita), so a municipality could effectively disguise the continuation of existing spending as the use of grant funds while allocating grant dollars according to local preferences. Depending on local preferences, municipalities could spend inframarginal funds entirely on sewerage capital (full pass-through), entirely on other goods and services (full crowd-out), or a mix of sewerage capital and other goods and services.

CWA grants were matching grants. Grants covered 75% of wastewater facility project costs, while municipalities covered the remaining 25% themselves.[10] This local cost-sharing does not guarantee full pass-

[7] In many cases, grant programs do not affect reduced form outcomes of interest unless they first cause an increase in spending. Gordon (2004) demonstrates this principle in the context of K-12 education by showing that previous findings of null effects of Title I transfers on student achievement can be explained by local governments reducing their own contribution to education spending in response to grant receipt.

[8] While this pattern often holds, estimates of grant pass-through are not consistent across settings (Hines and Thaler, 1995), and pass-through estimates can vary widely among similar grant programs. For example, Card and Payne (2002) studies the effects of school finance reforms between 1970 and 1992 on education spending and finds that a one dollar increase in state aid leads to a 60 cent increase in district education spending. Lutz (2010) examines a similar policy where state governments distributed unrestricted lump sum transfers for education to municipalities based on per-pupil property wealth and finds that these transfers are almost completely crowded out by changes in local spending.

[9] For example, researchers have documented the flypaper effect in intergovernmental grants for education (McGuire, 1978; Craig and Inman, 1986; Fisher and Papke, 2000; Card and Payne, 2002; Brunner et al., 2020), infrastructure (Gamkhar, 2000, 2003; Leduc and Wilson, 2017), welfare programs (Grossman and Roberts, 1989; Gamkhar and Oates, 1996) and police funding (Mello, 2019).

[10] In 1981, the federal share of costs was reduced to 55%.

through, or even at least 25% pass-through, but rather increases the pre-grant spending necessary for grant funds to be considered inframarginal.[11] In particular, a 1:3 local-to-federal cost-sharing ratio implies that existing spending would need to be larger than $\frac{4}{3}$ times the grant amount for funds to be completely inframarginal.[12]

The interaction between the Clean Water Act's grant program and wastewater treatment technology mandate yields several predictions for our grant pass-through estimates. The Act's mandate on wastewater treatment technology has the potential to ensure full pass-through when grant amount is set equal to upgrade costs, but inaccurate grant sizing allows for crowd-out (when grants are too big) or imposes a fiscal burden on grant recipients (when grants are too small). The mandate imposes no constraint on how municipalities use inframarginal grant funds for compliant facilities.[13] Consequently, for cities already in compliance with the mandate, we predict pass-through near zero; for cities constrained by the mandate, we expect pass-through to decline in grant size relative to estimated costs.[14]

Finally, the contemporary political climate toward public utilities informs how municipalities may have redistributed grant funds that were not bound by the mandate. When the CWA came into effect, public utilities were facing pressure to reduce user fees and become more competitive (Daigger, 1998). Since local governments had the means to redistribute grants to residents with little friction by adjusting utility prices, CWA grants to compliant municipalities could crowd out sewerage capital spending already in place relatively easily.[15] Compliant municipalities could then use this crowded out money to offset water and sewerage utility costs usually funded through water bills.

1.1.2.1 Target Parameter

The term "grant pass-through" can describe a number of target parameters. We focus on the increase in targeted spending induced by the average grant dollar.[16] This parameter, which we refer to as the "total pass-through rate", is equivalent to the increase in *total* targeted spending in *all* receiving municipalities relative to the *total* grant dollars distributed to *all* receiving municipalities. While this average can mask substantial

[11]Cost-sharing yields two different possible interpretations of pass-through. We define full pass-through as \$1 in grant funding leading to a \$1 increase in sewerage capital spending. Alternatively, since municipalities are required to pay 25% of project costs, one could define full pass-through as \$1 in grant funding leading to a \$1.33 increase in sewerage capital spending (the actual \$1 in grant aid plus the \$0.33 in local obligations). Our estimates are based on the first definition, appealing to the notion of federal funding itself being spent as desired. We present pass-through estimates using the second definition in Appendix Section 1.6.1.2.

[12]The median baseline spending in our sample (\$20 per capita) is approximately $\frac{4}{3}$ times the median yearly grant size (\$15 per capita).

[13]Compliant municipalities could still spend grant dollars on sewerage capital, but if they used this money to pay for upgrades that they already planned on making under their pre-grant budget constraint, CWA grants would not necessarily cause an increase in spending.

[14]Table 1.10 shows that compliant municipalities, non-compliant municipalities that received large grants and non-compliant municipalities that received small grants were observably similar before treatment. Table 1.11 shows that observable characteristics from 1971 do not predict grant size or compliance with the CWA's treatment technology mandate.

[15]The CWA instructed the EPA to issue guidelines to local governments on how to charge industrial and non-industrial users for waste treatment services, but did not place restrictions on adjusting rates.

[16]This is often the most policy-relevant parameter, but we may be interested in other parameters. For example, we may be interested in the increase in spending in the *average municipality* rather than the increase in expenditure induced by the *average dollar*. We estimate this "average pass-through rate" in Appendix Section 1.6.1.3.

heterogeneity, it allows for an accounting of the effectiveness of an entire grant program. It also has an interpretation that is convenient for conducting benefit to cost analyses; dividing the benefit to cost ratio of all grant dollars by the total pass-through rate will return an estimate of the cost-effectiveness of grant dollars that receiving municipalities actually spent on the targeted category of expenditure. Our research design will aim to recover the total pass-through rate of CWA grants.

1.2 Data

1.2.1 Clean Water Act Data

We obtain data on every CWA grant from the EPA's Grant Information Control System. This data contains information on the year that the EPA distributed each grant, which municipality received the grant, the specific wastewater treatment facility the grant was designated for and the amount distributed.[17]

Using a unique facility code, we merge our grant data to the 1972 Clean Watershed Needs Survey (CWNS), an assessment of the capital investment that publicly-owned wastewater treatment facilities required to come into compliance with the Clean Water Act. Importantly, the EPA conducted the 1972 CWNS before distributing any CWA grants, so the CWNS provides information on the treatment technology that each facility had in place before grant receipt.

The CWNS also includes information on the total amount of wastewater flowing through many facilities. We use these flow rates to calculate the costs that each non-compliant facility faced to upgrade its wastewater treatment technology using cost estimates from USEPA (1973), a separate EPA analysis of upgrade costs.[18]

1.2.2 Municipal Finance Data

We merge this linked facility level data to spending data from the municipality that operates each facility, dropping facilities operated by municipalities that we do not have spending data for. We refer to municipalities with only compliant facilities as compliant municipalities and those with only non-compliant facilities as non-compliant municipalities. Since we cannot observe exactly how local governments spend grants for specific treatment plants, we drop the 45 municipalities in our finance data with both compliant and non-compliant facilities from our sample.[19] Many municipalities receive multiple grants during our study period, so we define treatment as an absorbing state that begins when a municipality receives its first CWA grant.

Our data on municipal finances comes from the Census Bureau's Historical Database on Individual Government Finances. This database includes annual financial data starting in 1951 sourced from the *Com-*

[17]Our analysis does not include grants distributed through predecessor programs similar to the CWA. See Appendix Section 1.6.4.2 for further discussion.

[18]We do not calculate cost estimates for the 23 non-compliant municipalities in our sample that operate facilities that are missing information on total wastewater flow.

[19]Some facilities in the CWNS are missing information on compliance with treatment technology standards, however, no facilities operated by municipalities in our spending data are missing this information.

pendium of City Government Finances and *City Government Finances* surveys. Initially the Census Bureau surveyed the universe of municipalities with a population of at least 25,000. This limit rose to 50,000 in 1960 and 75,000 in 1987. As a result, we observe spending in 216 municipalities every year from 1951 until 1999.[20]

Table 1.1: 1970 Summary Statistics by Facility Compliance

	(1) Compliant	(2) Non-Compliant	(3) Difference
Population	115,501	396,525	-281,023
	(73,126)	(954,559)	(204,352)
Total Revenue	253.13	294.21	-41.08
	(138.17)	(155.15)	(36.16)
Total IGR	52.84	64.39	-11.55
	(62.93)	(70.45)	(16.43)
Federal IGR	11.43	10.85	0.58
	(15.64)	(11.56)	(2.96)
Nonwater Utility Rev	21.18	18.15	3.03
	(44.79)	(45.12)	(10.72)
Property Tax Revenue	82.21	98.30	-16.09
	(72.75)	(79.08)	(18.53)
Water Utility Revenue	23.10	22.91	0.19
	(9.10)	(8.58)	(2.06)
Total Expenditure	277.26	308.16	-30.90
	(154.56)	(166.63)	(39.10)
Total IG Exp	1.93	4.53	-2.60
	(4.54)	(9.98)	(2.19)
Capital Outlays	79.35	63.88	15.47
	(57.33)	(37.65)	(10.02)
Highway Exp	21.16	21.95	-0.79
	(8.38)	(9.41)	(2.19)
Sewerage Capital Outlay	8.94	6.17	2.77
	(9.61)	(6.80)	(1.76)
Sewerage Non-capital Exp	5.50	5.04	0.46
	(3.13)	(3.11)	(0.74)
Nonwater Utiltiy Exp	22.46	18.82	3.64
	(49.25)	(46.86)	(11.26)
Water IG Exp	1.98	2.39	-0.40
	(1.91)	(2.59)	(0.59)
Water Non-capital Exp	13.34	12.85	0.49
	(4.74)	(5.90)	(1.36)
Water Capital Outlay	13.45	8.37	5.08
	(19.01)	(11.53)	(3.16)
Observations	22	90	112

Notes: This table presents summary statistics for municipalities with compliant and non-compliant facilities in 1970, two years prior to the CWA. All covariates aside from the facility compliance dummy and population represent per capita values.

[20] Starting in 1967, the Census also collected spending data from a larger sample of municipalities once every five years. To obtain accurate pass-through estimates, we need to measure spending in all post-treatment periods, so we cannot use this data to estimate pass-through. We re-estimate the crowd-out results from Section 1.4 using this data in Appendix 1.6.2.

Variables concerning municipal spending on water and sewerage utilities are available beginning in 1956. The data separates expenditure toward sewerage services into capital outlays and non-capital expenditure. We also observe revenue raised through a city's water and sewerage utility services (i.e. water bills). Because timing of grant receipt varies, we adjust both grant amount and all municipal spending outcomes to 1973 dollars.

Figure 1.2: Map of Municipalities

Notes: This figure plots the locations of the 112 municipalities in our sample. Non-compliant municipalities are shown as Xs and compliant municipalities are shown as points.

States had some discretion about where they distributed grants, so grant receipt might be correlated with trends in municipal spending. Rather than leveraging variation in grant *receipt*, we focus our analysis on the municipalities in our finance data that received at least one CWA grant and estimate pass-through by leveraging variation in grant *timing*.[21] Table 1.1 presents summary statistics for these 112 municipalities from 1970, two years before Congress passed the CWA, split based on whether or not the city's wastewater treatment facilities were in compliance with the CWA's capital mandate when the Act came into effect.[22] All variables besides population are in per capita terms. Compliant and non-compliant municipalities demonstrate no statistically distinguishable differences across any of the observable characteristics presented in Table 1.1, though this is largely due to statistical imprecision from the small sample size. Comparisons of means suggest that, on average, non-compliant municipalities in our sample had larger populations and greater total revenues and expenditures than compliant municipalities. Taken at face value, transfer revenue from other governments to non-compliant municipalities and more lucrative property taxes explain this difference. Figure 1.2 shows

[21] We obtain similar results if we use the municipalities in our finance data that never received a CWA grant as a control group in Appendix Section 1.6.1.4.

[22] We show that observable spending characteristics do not predict grant timing in Appendix Section 1.6.1.5.

the location of each municipality in our sample.[23]

1.3 Grant Pass-Through

1.3.1 Pass-Through Methods

We want to select a design that recovers the total pass-through rate. To do this, we need an estimate of the average treatment effect on the treated (ATT), which describes the average increase in sewerage capital expenditure after grant receipt. Dividing the ATT by average grant amount returns the pass-through rate for the average CWA grant dollar in our sample.[24]

We follow Callaway and Sant'Anna (2020b) and leverage variation in the timing of CWA grant receipt to estimate "group-time average treatment effects". At least one municipality receives its first CWA grant in every year from 1972 to 1981, as well as in 1985 and 1988. This variation in treatment timing yields 12 "timing groups", which we index $g = 1972, ..., 1981, 1985, 1988$. We estimate the group-time average treatment effect for each group g in each time period t by comparing units in g to units that were not yet treated in time t. We then summarize the reduced form effect of grant receipt on sewerage capital spending by aggregating these group-time average treatment effects together.[25]

Under a parallel trends assumption, the group-time average treatment effect for group g at time t is

$$ATT(g,t) = E[C_t - C_{g-1} | G_g = 1] - E[C_t - C_{g-1} | D_t = 0]$$

where C_t is average sewerage capital spending at time t, G_g is a dummy variable that equals one for units in timing group g and D_t is a dummy that equals zero for units not-yet-treated at time t. We estimate each $ATT(g,t)$ with its sample analogue, $\widehat{ATT(g,t)}$.[26] This process yields many $\widehat{ATT(g,t)}$, most of which are identified off of relatively few observations, so instead of interpreting individual $\widehat{ATT(g,t)}$, we summarize the effect of treatment by aggregating the $\widehat{ATT(g,t)}$ together.

[23] See Appendix 1.6.4 for more discussion of our data.

[24] Two-way fixed effects (TWFE) estimators are one of the most common ways to leverage variation in treatment timing. This estimator regresses targeted spending on a treatment dummy and unit and time fixed effects. The resulting estimate is a weighted average of comparisons between (1) spending in newly treated municipalities relative to spending in municipalities that have not yet been treated and (2) spending in newly treated municipalities relative to spending in already-treated municipalities. While the first type of comparison only requires a parallel trends assumption to recover an estimate of the ATT, to use the second type of comparisons, we must assume that treatment effects are constant over time. If treatment effects are dynamic, already-treated units, which are still actively responding to treatment, will not make valid counterfactuals for what would have happened in the absence of treatment, so the second type of comparison does not return a meaningful result and we cannot interpret a TWFE coefficient as the ATT (Goodman-Bacon, 2021a). We have an a priori reason to expect the effect of grant receipt on spending to change over time; states distributed CWA grants to fund specific projects, so we would expect grant receipt to cause an increase in sewerage capital spending that only lasts until the project is completed. When spending returns to pre-treatment levels, this decrease is subtracted from a TWFE estimate, biasing the estimate upwards. Since this estimator does not return the ATT, we cannot recover the total pass-through rate from a TWFE estimate, which motivates us to use a different estimator.

[25] Since this estimator only relies on comparisons between newly treated units and not-yet-treated units, we do not need to place any restrictions on treated potential outcomes, and our estimates will not be biased by dynamic treatment effects.

[26] $\widehat{ATT(g,t)} = E_n[\widehat{C}_t - \widehat{C}_{g-1} | G_g = 1] - E_n[\widehat{C}_t - \widehat{C}_{g-1} | D_t = 0]$.

We aggregate group-time treatment effects into summary measures in several ways. First, we examine how the effect of CWA grant receipt on sewerage capital expenditure evolves over time with an event study specification. For each time relative to treatment e, we estimate the effect of treatment for units that have been treated for e periods with $\widehat{\theta}_D(e)$.

$$\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e) \tag{1.1}$$

Taking an average of the $\widehat{\theta}_D(e)$ for e such that $e \geq 0$ provides a summary measure of the reduced form effect of grant receipt on sewerage capital spending.

Another way to summarize this effect is to take a weighted average of all of the $\widehat{ATT(g,t)}$ such that $g \leq t$ where the weights are based on timing group size. The formula for this weighted average is

$$\frac{1}{\kappa} \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\}\widehat{ATT(g,t)}P(G=g) \tag{1.2}$$

where $\kappa = \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\}P(G=g)$.

We can use these estimates to recover the total pass-through rate. Under a parallel trends assumption, our summary measures of the effect of grant receipt on sewerage capital spending represent estimates of the average increase in targeted spending after grant receipt. We can repeat our estimation process with grant amount as the dependent variable to estimate the associated first stage. Dividing the reduced form effect of grant receipt on sewerage capital spending by the first stage relationship between grant receipt and grant amount returns the average increase in sewerage capital spending induced by a dollar of grant funding, which is the total pass-through rate.

To identify a given $ATT(g,t)$, we must assume that, in absence of treatment, the average outcomes in timing group g would have followed parallel trends with not-yet-treated groups in period t. Since we are primarily interested in summary measures, we do not need to interpret individual $\widehat{ATT(g,t)}$ and we do not need this assumption to hold for all group-time pairs. Instead, we must assume that, on average, parallel trends holds for all groups g in all periods t such that $g \leq t$.

Callaway and Sant'Anna (2020a) constructs standard errors with a multiplier bootstrap procedure. Instead of re-sampling observations as in a pair-bootstrap, each multiplier bootstrap draw perturbs the influence function of the estimate (which measures the dependence of the estimate on each cluster in the sample). Our pass-through estimator is a function of two Callaway and Sant'Anna (2020a) objects (specifically, the reduced for divided by the first stage), and rather than derive the influence function for this estimator, we present pair-bootstrap standard errors clustered at the municipality level for our pass-through estimates.

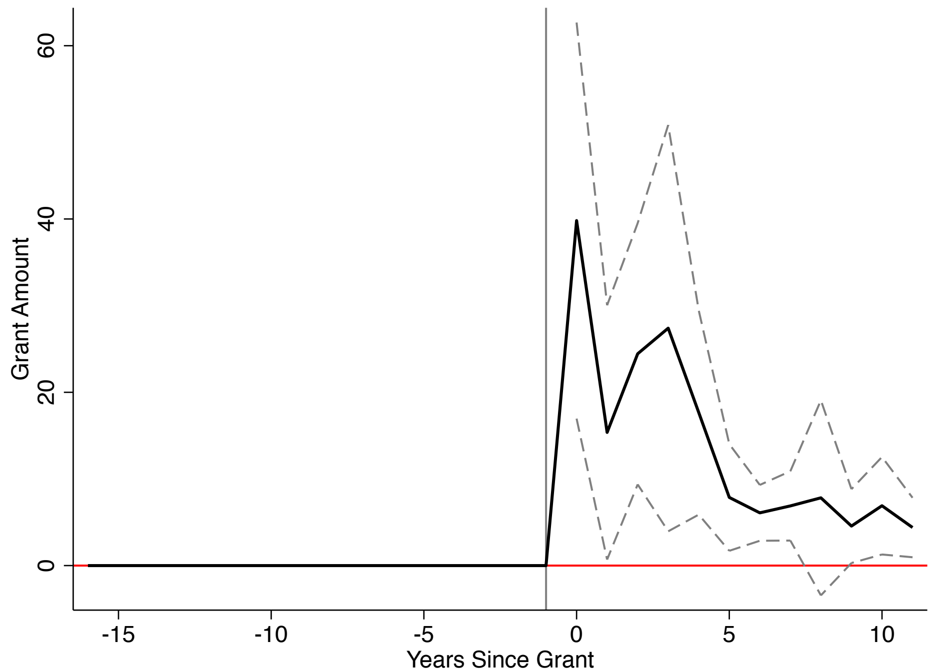


Figure 1.3: First Stage Relationship Between Grant Receipt and Grant Amount

Notes: This figure shows average grant amount for each year relative to treatment.

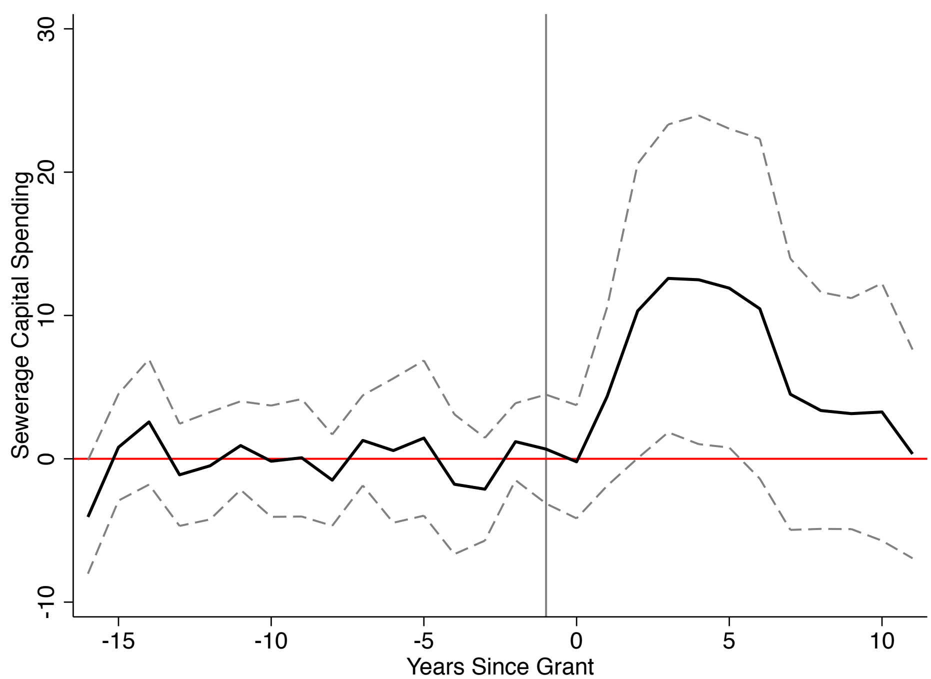


Figure 1.4: Sewerage Capital Spending Increased After Grant Receipt

Notes: This figure shows the reduced form relationship between grant receipt and sewerage capital spending by presenting the $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\} \widehat{ATT(g,t)} P(G=g|t-g=e)$. g indexes timing groups and t indexes years. $\widehat{ATT(g,t)}$ is the difference between changes in spending in units in group g in periods t and $g-1$ and changes in spending in units not yet treated at time t in periods t and $g-1$.

Table 1.2: CWA Grants had Low Pass-Through

	(1)	(2)	(3)
Panel A: Dynamic Aggregation	Grant Amount	Sewerage Capital	
Treat * Post	11.72***	5.628***	
	(1.31)	(1.746)	
Pass-through			0.480**
			(0.207)
p-value: Pass-through = 0			0.020
p-value: Pass-through = 1			0.012
Observations	3630	3630	3630
Panel B: Dynamic Aggregation ($e \leq 6$)			
Treat * Post	19.81***	8.844***	
	(2.768)	(2.485)	
Pass-through			0.446***
			(0.142)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630
Panel C: Balanced Dynamic Aggregation			
Treat * Post	19.87***	8.914***	
	(3.011)	(2.302)	
Pass-through			0.448***
			(0.137)
p-value: Pass-through = 0			0.001
p-value: Pass-through = 1			0.000
Observations	3564	3564	3564
Panel D: Simple Aggregation			
Treat * Post	13.26***	6.115***	
	(1.526)	(1.980)	
Pass-through			0.461***
			(0.156)
p-value: Pass-through = 0			0.003
p-value: Pass-through = 1			0.001
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of the first stage relationship between grant receipt and grant amount, the reduced form effect of grant receipt on sewerage capital spending and the pass-through rate for all CWA grants. Panel A presents averages of the $\widehat{\theta}_D(e)$ for each time relative to treatment e for $e \geq 0$ calculated with $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\} \widehat{ATT(g,t)} P(G=g|t-g=e)$. Panel B re-estimates the results in Panel A for e such that $0 \leq e \leq 6$. Panel C re-estimates the results in Panel B for municipalities that we observe for at least seven post-treatment periods. Panel D presents averages of group-time treatment effects based on group size calculated with the following equation $\frac{1}{\kappa} \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\} \widehat{ATT(g,t)} P(G=g)$ where $\kappa = \sum_{g=1972}^{1985} \sum_{t=1972}^{1987} 1\{g \leq t\} P(G=g)$. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in column 2. Column 3 shows the implied pass-through rate from dividing the reduced form in column 2 by the first stage in column 1.

1.3.2 Full Sample Pass-Through Results

We begin with a flexible specification that explores how the effect of grants evolves over time. Figure 1.3 examines the first stage relationship between grant receipt and grant amount by presenting the $\widehat{\theta}_D(e)$ from equation 1.1 with grant amount as the dependent variable. Each $\widehat{\theta}_D(e)$ in Figure 1.3 is equal to average grant amount e years after treatment. Figure 1.4 then presents the associated reduced form estimates of the effect of grant receipt on sewerage capital spending. The null effects in the 16 years before grant receipt support a research design that leverages variation in grant timing by showing common trends in sewerage capital spending in treated and not-yet-treated municipalities before treatment, which suggests that parallel trends would have continued in absence of treatment. The estimates increase after the arrival of a municipality's first CWA grant and remain high for seven years after grant receipt. Sewerage capital spending decreases to near pre-treatment levels by eight years after grant receipt. The EPA estimates that upgrades paid for with CWA grants could take up to 10 years from initial grant receipt to project completion (USEPA, 2002), so the shapes of Figures 1.3 and 1.4 are consistent with grant funding and sewerage capital spending returning to pre-treatment levels as municipalities complete grant funded projects.[27]

Table 1.2 summarizes the results in Figures 1.3 and 1.4. First, Panel A presents averages of the $\widehat{\theta}_D(e)$ from Figures 1.3 and 1.4. Since the effect of grant receipt on both sewerage capital spending and grant amount return to near pre-treatment levels by seven years after treatment, Panel B re-calculates the results in Panel A for the fist seven post-treatment periods.

Since treatment timing varies, we do not observe all of the municipalities in our sample for the same number of post-treatment periods. For this reason, the composition of the groups that contribute to each event study coefficient in Figures 1.3 and 1.4 changes when event time changes. This causes our dynamic aggregations in Panels A and B of Table 1.2 to weight up observations of municipalities that are treated earlier, which can produce misleading results if the effect of grant receipt varies across timing groups. We address this in Panel C by re-calculating the results in Panel B on the sub-sample of municipalities that we observe for at least seven post-treatment periods. The estimates are similar to those in Panel B.[28]

Finally, Panel D of Table 1.2 presents a weighted average of all post-treatment group-time-treatment effects calculated with equation 1.2.

Dividing the reduced form estimates in Table 1.2 by their respective first stage estimates suggests that CWA grants have a total pass-through rate between 0.446 and 0.480. Across aggregations, the 95% confidence intervals of the first stage and reduced form estimates do not overlap and we can reject a test of the hypothesis that our pass-through estimate equals one, so we reject full grant pass-through for the full sample.

[27]Table 1.12 presents the results in Figures 1.3 and 1.4 in tabular form.

[28]Figure 1.10 presents the associated event study.

1.3.3 Evaluating CWA Investments

We can use estimates of the total pass-through rate for the full sample to find the benefit to cost ratio of the CWA grant dollars that municipalities spent on sewerage capital (i.e., the benefit to cost ratio of the wastewater capital upgrades themselves). Using increased housing prices to quantify the benefits of CWA grants, Keiser and Shapiro (2019a) estimates a benefit to cost ratio of 0.26. Flynn and Marcus (2021) finds that this ratio may be as high as 0.45 after incorporating improvements to infant health.[29] Dividing this benefit to cost ratio by the total pass-through rate yields the benefit to cost ratio of CWA grant dollars that municipalities spent on sewerage capital. Our preferred specification in Panel B of Table 1.2 shows that, in total, receiving governments in our sample spent 44.6% of CWA grant dollars on sewerage capital upgrades. Dividing these results implies that the CWA grant dollars that municipalities spent on sewerage capital have a benefit/cost ratio of 1.01.[30] This suggests that, given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded wastewater capital upgrades.[31]

The standard error on our preferred specification suggests a 95% confidence interval of 0.165 to 0.727. Dividing previously estimated benefits of CWA grants by the upper bound lets us reject a benefit/cost ratio less than 0.62.

We examine which municipalities are driving this low pass-through in the next section.

1.3.4 Heterogeneity in Pass-Through

We explore heterogeneity in pass-through by mandated costs. Compliant municipalities face no mandated costs, and mandated costs vary within our sample of non-compliant municipalities. Since, in theory, setting grant amount equal to mandated costs increases the likelihood of full pass-through, we estimate pass-through separately for compliant municipalities, non-compliant municipalities where grant amount is approximately equal to mandated costs, and non-compliant municipalities where grant amount is greater than mandated costs.

Splitting the sample along compliance is straightforward, but it is not immediately obvious how to divide non-compliant municipalities based on mandated costs. We first need to identify which grants are approximately equal to mandated costs and which grants are larger than mandated costs. Once we define these groups, we can separately estimate pass-through for (1) non-compliant municipalities whose grant funding is completely bound by the capital mandate and (2) non-compliant municipalities that receive grants that include funding that is not bound by the capital mandate. Figure 1.1 shows that, while grants are similar to

[29]This assumes that hedonic estimates do not capture any health benefits. It is unlikely that this historical population fully understood the relationship between surface water quality and the health of infants in utero, so this assumption is plausible.

[30]The other results in Table 1.2 suggest that the CWA grant dollars that municipalities spent on sewerage capital had a benefit/cost ratio of between 0.937 and 1.01. This benefit/cost ratio is even higher if we use the estimates from our stacked difference-in-difference in Table 1.24, which suggest a benefit/cost ratio of 1.32.

[31]We discuss potential shortcomings of this back-of-the-envelope calculation in Appendix Section 1.6.1.7.

costs for municipalities that receive relatively small grants, grants to municipalities that receive large grants are substantially larger than the estimated cost of upgrading. For this reason, we want to divide our sample of non-compliant municipalities into those that received large and small grants.

We divide our sample at the point on the distribution of grant amount where costs are no longer increasing in grant size.[32] To estimate this point, we keep one observation from each non-compliant municipality that we have cost data for and estimate the following equation with non-linear least squares

$$\begin{aligned} cost_i =& (a_1 + b_1 * TotalGrant_i) * 1\{TotalGrant_i < split\} \\ & + (a_1 + b_1 * split + b_2 * (TotalGrant_i - split)) * 1\{TotalGrant_i >= split\} \end{aligned} \tag{1.3}$$

where $cost_i$ is our estimate of municipality i's total per capita cost of upgrading and $TotalGrant_i$ is municipality i's total per capita grant amount. Column 1 of Table 1.3 presents our estimate of $split$. This result shows that costs are increasing in grant size up to \$125 of total per capita in grants and flatten out for grants above \$125. We estimate pass-through separately for non-compliant municipalities above and below this threshold.[33]

Table 1.3: Non-Linear Least Squares Estimates of Break in Cost

	(1) Total Cost	(2) Total Cost
Split	125.2**	138.2
	(60.49)	(106.8)
Observations	41	58

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents $split$ from estimating $cost_i = (a_1 + b_1 * grant_i) * 1\{grant_i < split\} + (a_1 + b_1 * split + b_2 * (grant_i - split)) * 1\{grant_i >= split\}$ on a sample of non-compliant municipalities with non-linear least squares. Column 1 presents results from estimating this equation with total cost calculated from estimates in USEPA (1973) as the dependent variable. Column 2 re-estimates the result in column 1 without dropping municipalities in the bottom decile of total grant amount.

1.3.4.1 Semi-Parametric Methods

Rather than imposing parametric restrictions on the relationship between grant amount and targeted spending, we explore heterogeneity in pass-through using an approximation of the dose-response function relating grant amount and sewerage capital spending. As shown in Figure 1.1, there is substantial variation in grant amount,

[32] This is the point where the quadratic fit in Figure 1.8 flattens out.

[33] The EPA did not include facilities with a total flow rate of less than 5 million gallons per day in their sample when estimating the cost functions that we use to calculate costs. These cost functions over-estimate costs for smaller facilities (USEPA, 1973). Most municipalities in the bottom decile of grant aid had facilities of this size, so we drop these municipalities when we estimate the results in column 1 of Table 1.3. We re-estimate equation 1.3 on the full sample in column 2, which yields an estimated cutoff of \$138 per capita. Table 1.13 shows that our pass-through estimates are similar if we use either of the cutoffs in Table 1.3.

and plotting out each municipality's total grant amount will give us the x-axis of this dose-response function. We can then find the y-axis by estimating each municipality's response to grant receipt.

To do this, we follow Deshpande and Li (2019) and re-organize our data into "stacks". Each stack S is defined by a single treated municipality, called municipality s, which is labeled as treated in that stack. We then add municipalities that received grants at least seven years after the treated municipality to the stack, which are labeled as controls. We can estimate the reduced form effect of grant receipt on sewerage capital spending for municipality s by comparing spending between municipality s and the control municipalities in stack S before and after the year in which municipality s becomes treated.[34] Estimating equation 1.4 on observations in stack S returns the result from this comparison.

$$C_{it} = \alpha_0 + \beta_s^{2X2} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it} \tag{1.4}$$

Municipalities are indexed by i, stacks by s, and years by t. C_{it} is per capita sewerage capital spending, and $treat_i$ is a dummy variable that equals one for municipality s (the treated unit). Plotting β_s^{2X2} against the total amount of grant dollars distributed to municipality s for all stacks yields a semi-parametric approximation of the dose-response relationship between grant amount and sewerage capital spending.[35]

We summarize these figures by appending groups of stacks into one dataset and estimating a stacked difference-in-difference. Equation 1.5 estimates first stage relationship between grant receipt and grant amount, denoted g_{it},

$$g_{it} = \alpha_0 + \beta_{fs}^{stacked} D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{1.5}$$

and equation 1.6 estimates the reduced form effect of grant receipt on sewerage capital spending,

$$C_{it} = \alpha_0 + \beta_{rf}^{stacked} D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{1.6}$$

where D_{it} is a dummy variable that equals one after grant receipt, and α_{is} and α_{ts} are stack-by-municipality

[34] By estimating the effect of grant receipt on *individual units* across *all post treatment periods*, a stacked estimator allows us to semi-parametrically examine the relationship between grant amount and sewerage capital spending before constructing summary measures of pass-through. Since timing groups contain both compliant and non-compliant municipalities, as well as municipalities that received different sized grants, we cannot repeat this process with Callaway and Sant'Anna (2020a) without making a priori assumptions about the relationship between grant amount and sewerage capital spending, as well as the relationship between compliance and sewerage capital spending.

[35] To identify the β_s^{2X2}, we must assume that sewerage capital spending in municipality s and the control municipalities in stack S would have followed parallel trends in absence of treatment. This is stronger than the assumption we need to identify $\beta_{rf}^{stacked}$ in equation 1.6, which only requires us to assume that parallel trends holds on average (we derive this identifying assumption in Appendix Section 1.6.3.4). For this reason, we should interpret individual β_s^{2X2} with caution. Note that, under parallel trends, each β_s^{2X2} represents the effect of going from zero grant dollars to the grant dollars distributed to municipality *s on municipality s*. This does not tell us anything about the effect of going from municipality $s-1$'s grant amount to municipality s's grant amount, or the effect of grant receipt on any other municipality.

and stack-by-year fixed effects. In tandem, theses fixed effects force identification to come from within-stack variation, ensuring that our estimates will not reflect any comparisons between newly treated municipalities relative and already-treated municipalities. Since it is only identified off of within-stack comparisons, $\beta_{rf}^{stacked}$ is an unweighted average of all the β_s^{2X2}. We present pair-bootstrap standard errors clustered at the municipality level.[36]

1.3.4.2 Pass-Through Results for Sub-Groups of CWA Grants

Figure 1.5 examines the relationship between grant amount and sewerage capital spending in non-compliant municipalities. This figure plots out the β_s^{2X2} from estimating equation 1.4 on units in each non-compliant stack against municipality s's total grant amount.[37] Sewerage capital spending is increasing in grant amount for municipalities that received grants totaling up to \$125 per capita, but this relationship is flat for municipalities receiving more than \$125 per capita. Importantly, Figure 1.5 shows a relationship between the β_s^{2X2} and total grant amount that is very similar to the relationship between cost and total grant amount shown in Figure 1.1.

Figure 1.6 repeats this process for compliant municipalities. This figure shows that grants to compliant municipalities had low pass-through regardless of size.

Table 1.4 summarizes Figures 1.5 and 1.6 by estimating equations 1.5 and 1.6 on sub-samples of stacks. Panel A presents pass-through estimates for non-compliant municipalities that received grants totaling up to \$125 per capita. The pass-through estimate in column 3 suggests that CWA grants led to a dollar-for-dollar increase in sewerage capital expenditure in these municipalities. Panel B presents estimates for all other stacks defined by non-compliant municipalities, which have much lower pass-through. Panel C presents estimates for all stacks defined by compliant municipalities, which also have low pass-through.

Interpreting the results in Table 1.4 as the total pass-through rate suggests that non-compliant municipalities in our sample that received grants totalling less than \$125 dollars per capita increased sewerage capital spending by 104.9% of all grant dollars distributed to them, while all other non-compliant municipalities in our sample increased sewerage capital spending by 26.7% of all grant dollars distributed to them. Compliant municipalities in our sample increased sewerage capital spending by 3.37% of grant dollars distributed to them.[38]

[36]To construct standard errors, we take a random sample (with replacement) of municipalities in the unstacked data, re-form stacks, then perform our estimates. We repeat this process 1000 times and use the results to calculate bootstrap standard errors. See Appendix Section 1.6.3.4 for further discussion of stack construction and inference with a stacked difference-in-differences estimator.

[37]Because they reflect averages across seven years, we multiply the β_s^{2X2} in Figures 1.5 and 1.6 by seven. This makes these figures more comparable in scale to Figure 1.1 and with one another.

[38]Table 1.14 shows that pass-through was low for compliant municipalities that received large and small grants alike, which suggests that the heterogeneity in pass-through by grant size documented in Panels A and B of Table 1.4 is driven by differences in costs rather than other confounding factors associated with grant size that are similar across compliant and non-compliant municipalities.

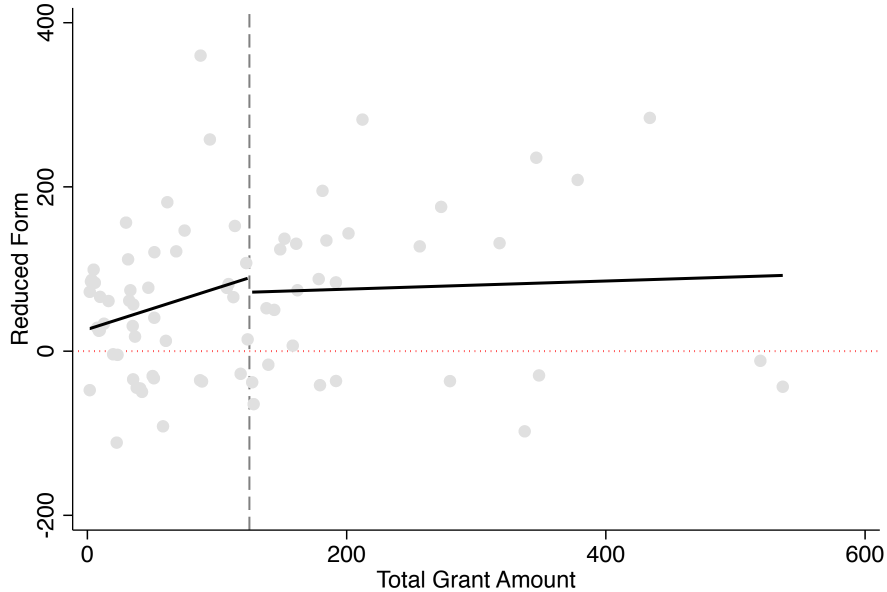


Figure 1.5: 2X2 DD Coefficients vs Grant Amount for Non-Compliant Municipalities

Notes: This figure plots the β_s^{2X2} from estimating $C_{it} = \alpha_0 + \beta_s^{2X2} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it}$ on observations in each stack s (for stacks defined by non-compliant municipalities). Sewerage capital spending is the dependent variable. We plot each stack's β_s^{2X2} against the total grant dollars distributed to the treated municipality in that stack. We divide the stacks at \$125 per capita of total grant dollars and fit lines to the β_s^{2X2} in each group. We multiply the β_s^{2X2} by seven to reflect that each β_s^{2X2} represents an average across seven post-treatment years.

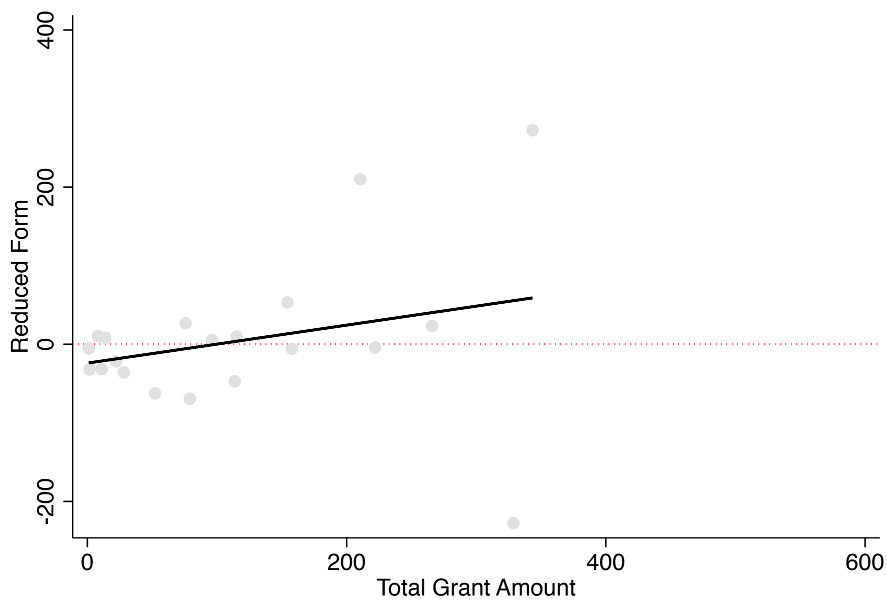


Figure 1.6: 2X2 DD Coefficients vs Grant Amount for Compliant Municipalities

Notes: This figure re-estimates the results in Figure 1.5 for stacks defined by compliant municipalities.

Table 1.4: Pass-Through is Heterogeneous in Grant Size and Compliance

	(1)	(2)	(3)
Panel A: Non-Compliant and < \$125	Grant Amount	Sewerage Capital	
Treat * Post	6.935***	7.279	
	(0.729)	(7.653)	
Pass-through			1.049
			(1.138)
p-value: Pass-through = 0			0.357
p-value: Pass-through = 1			0.965
Observations	2162	2162	2162
Panel B: Non-Compliant and >= \$125			
Treat * Post	46.01***	11.81	
	(6.306)	(7.130)	
Pass-through			0.267
			(0.156)
p-value: Pass-through = 0			0.087
p-value: Pass-through = 1			0.000
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	16.43***	0.553	
	(3.905)	(4.655)	
Pass-through			0.0337
			(0.345)
p-value: Pass-through = 0			0.922
p-value: Pass-through = 1			0.009
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$. Municipalities are indexed by i, stacks by s, and years by t. D_{it} is a dummy variable that equals one after grant receipt, and α_{is} and α_{ts} are stack-by-municipality and stack-by-year fixed effects. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 uses D_{it} as a instrument for grant amount, which returns a coefficient equivalent to dividing the reduced form in column 2 by the first stage in column 1. Panel A presents estimates for stacks defined by non-compliant municipalities who received grants totalling less than \$125, Panel B for all other non-compliant stacks, and Panel C for all stacks defined by compliant municipalities. We present p values from testing for a pass-through rate of one at the bottom of each panel.

Based on the above results, we reject full pass-through for compliant municipalities and non-compliant municipalities receiving more than \$125 per capita in grant aid, but cannot reject full pass-through for non-compliant municipalities receiving grants totaling less than \$125 per capita, as expected.[39] While we cannot reject complete crowd-out in non-compliant municipalities that received grants totaling up to \$125 per capita or equality of the pass-through estimates in column 3 Table 1.4, these results, along with the shapes of Figures 1.5 and 1.6, are consistent with municipalities only increasing sewerage capital spending up to the point where

[39]We obtain similar results when we re-estimate the results in Table 1.4 with Callaway and Sant'Anna (2020a). We present these results in Appendix Section 1.6.1.6, and discuss our choice of estimators in Appendix 1.6.3.

they are in compliance with the CWA's capital mandate.[40]

1.3.5 Does Variation in Grant Size Recover Total Pass-Through?

Researchers often estimate pass-through by leveraging variation in grant size with a dose-response two-way fixed effects estimator. In this section, we follow Callaway et al. (2021) to show that, when the effect of a grant of a given size is not constant across receiving municipalities, or when the effect of grant receipt changes over time, this type of estimator may not return the total pass-through rate.

Consider, as an example, the pass-through estimates in Keiser and Shapiro (2019a), which suggest that municipalities that received any CWA grant spent \$0.94 of every federal dollar on sewerage capital by estimating

$$C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{1.7}$$

where C_{it} is cumulative sewerage capital spending from 1970 to 2002 and g_{it} is cumulative grant amount. Similar to a binary TWFE estimate identified off of variation in treatment timing, β_{dr}^{TWFE} is a weighted average of (1) comparisons of newly treated municipalities relative to municipalities that have not yet been treated, (2) comparisons of newly treated municipalities relative to already-treated municipalities, and (3) comparisons of municipalities that received large grants relative to municipalities that received smaller grants.

To interpret comparisons of newly treated municipalities relative to already-treated municipalities as the ATT, we need to make additional assumptions along with parallel trends. When treatment effects are dynamic, already-treated units, which are still actively responding to treatment, will not make valid counterfactuals for what would have happened in the absence of treatment. For this reason, we must assume that treatment effects do not change over time to use this type of variation (Goodman-Bacon, 2021a).

Identifying a meaningful effect using comparisons of municipalities that received large grants relative to municipalities that received smaller grants requires additional assumptions as well. In particular, we must assume that a grant of a given size will have the same effect on sewerage capital spending in all receiving municipalities in our sample (Callaway et al., 2021). If this assumption does not hold, municipalities that received small grants will not be valid counterfactuals for what would have happened if a municipality that received a large grant had received a small grant instead, and this type of comparison can either over- or under-estimate the total pass-through rate.

The results in Sections 1.3.2 and 1.3.4.2 suggest that the effect of CWA grants on sewerage capital spending is not homogeneous across time relative to treatment, or across receiving municipalities. As shown in

[40] We show that differences in pass-through between sub-groups are not driven by differences in baseline sewerage capital spending in Appendix Section 1.6.1.8.

Figure 1.4, treatment effects change over time as grant funding runs out. Additionally, Figure 1.5 suggests that the effect of grants to non-compliant municipalities is increasing in grant size (up to the point where the capital mandate is satisfied), while Figure 1.6 suggests that grants to compliant municipalities had little to no effect on sewerage capital spending regardless of grant size. For this reason, a grant of a given size would likely affect spending differently in a compliant municipality than in a non-compliant municipality, which violates the homogeneity assumption.[41]

The reliability of an estimate of β_{dr}^{TWFE} from equation 1.7 depends on how much of the estimate is identified off of comparisons of early treated units to not-yet-treated units. Table 1.5 decomposes β_{dr}^{TWFE} using Goodman-Bacon (2021b), which shows how much of β_{dr}^{TWFE} is identified off of comparisons of newly treated units to not-yet-treated units, comparisons of newly treated units to already treated units, and comparisons of units that received large grants to units that received smaller grants.[42]

Table 1.5: Decomposition of Keiser and Shapiro (2019a) Pass-Through Estimates

	(1) Weight	(2) Average Estimate
Newly Treated vs Not-Yet-Treated	0.017	0.428
Newly Treated vs Already-Treated	0.553	1.002
Large Grant vs Small Grant	0.430	0.835

Notes: This table presents the average difference-in-difference estimates and their weights from decomposing β_{dr}^{TWFE} from $C_{it} = \alpha_0 + \beta_{dr}^{TWFE} + \alpha_i + \alpha_t + \varepsilon_{it}$ using Goodman-Bacon (2021b). C_{it} is cumulative sewerage capital spending, g_{it} is cumulative grant amount, and α_i and α_t are municipality and year fixed effects.
Source: Keiser and Shapiro (2019a)

Row 1 of Table 1.5 describes the part of β_{dr}^{TWFE} identified off of comparisons of newly treated units to not-yet-treated municipalities. This variation suggests a pass-through rate of 0.428, which is similar to the pass-through estimates in Table 1.2.[43] This represents an estimate of the total pass-through rate, but less than two percent of β_{dr}^{TWFE} comes from this type of comparison.

Row 2 of Table 1.5 shows that more than half of β_{dr}^{TWFE} comes from comparisons of late treated municipalities to already-treated municipalities. Consistent with bias from treatment effects that decrease over time, comparisons of newly treated municipalities to already-treated municipalities yield a result that is more than

[41] Estimating β_{dr}^{TWFE} on compliant and non-compliant municipalities separately would solve the problem introduced by heterogeneity in treatment effects by compliance, but would not address problems caused by dynamic treatment effects. For this reason, estimating β_{dr}^{TWFE} on sub-groups of municipalities will not necessarily return the total pass-through rate for these sub-groups. See Appendix Section 1.6.3.2 for further discussion of the assumptions required to use this type of variation.

[42] This program is the analogue of Goodman-Bacon et al. (2019) for difference-in-difference estimates identified off of continuous variation.

[43] The part of the pass-through estimate in Keiser and Shapiro (2019a) that is identified off of comparisons of newly-treated units to not-yet-treated is identified off of the same type of variation as the pass-through estimates in Table 1.2. The similarity of these results suggests that, even though we make several sample restrictions that Keiser and Shapiro (2019a) do not (we drop both municipalities that are missing information on compliance and municipalities in our finance data with both compliant and non-compliant facilities), the differences between our estimates and those in Keiser and Shapiro (2019a) come from differences in the source of variation we use rather than differences between our samples.

twice as large as the result from comparing newly treated municipalities to not-yet-treated municipalities.

Finally, Row 3 of Table 1.5 shows that a large portion of β_{dr}^{TWFE} comes from comparisons of municipalities that received large grants to those that received smaller grants. Many of these comparisons are between compliant and non-compliant municipalities, which do not represent causal estimates.

The decomposition in Table 1.5 shows that, when the effect of a grant of a given size is heterogeneous across receiving municipalities, estimating pass-through with a dose-response TWFE estimator will produce an estimate that may not equal the total pass-through rate.

Table 1.6: Pass-Through Results are Robust to Population Weighting

	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	14.55***	5.628***	
	(3.45)	(1.69)	
Pass-through			0.387**
			(0.162)
p-value: Pass-through = 0			0.017
p-value: Pass-through = 1			.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results from Panel B of Table 1.2 weighting by population.

Table 1.7: Full Sample of Grants

Panel A: All Grants	(1) All	(2) Non-compliant	(3) Compliant
Total Grant Amount	2,741,595	2,454,547	2,260,387
Total Cost		932,683	
Observations	10,202	9027	940
Panel B: Grants with Spending Data			
Total Grant Amount	46,883,636	54,197,768	17,627,110
Total Cost		28,815,210	
Observations	112	90	22

Notes: Column 1 presents average total costs and average total grant amount for all municipalities that received CWA grants, and column 2 presents average total costs and average total grant amount the municipalities in our sample.

1.3.6 Can We Generalize These Results to All CWA Grants?

Our results represent the pass-through of \$25.5 billion (2014\$) of the \$153 billion in grants funded by the CWA, and in this section, we discuss how generalizable our results are to CWA grants outside of our sample.

The Census Bureau collects annual finance data from the universe of municipalities with populations over

75,000, so our sample only includes grants distributed to relatively large cities. If population is correlated with preferences for sewerage capital spending, grants to municipalities outside of our sample might have a different pass-trough rate than grants to municipalities within our sample. We adjust for heterogeneity in pass-through by population in Table 1.6 by weighting our pass-through results by population. This does not change our results in a meaningful way, which suggests that, within our sample, treatment effects are not correlated with population.[44] Since the Census data that we use is the only comprehensive source of state and local government finance data collected on a nationwide scale (US Census Bureau, 2021), we cannot test if this homogeneity holds *outside of our sample*, but the similar pass-through rates across municipality size *within our sample* suggests that CWA grant pass-through is not strongly correlated with population.

Since pass-through is correlated with mandated costs, our pass-through estimates might not be generalizable to all CWA grants if municipalities outside of our sample faced different mandated costs than municipalities in our sample. We compare municipalities in our sample to municipalities outside of our sample in Table 1.7. Municipalities outside of our sample received smaller grants than those in our sample, but they also faced lower mandated costs, so the ratio of grant size to mandated costs in municipalities within and outside of our sample is relatively similar. It is not unreasonable to expect non-compliant municipalities outside of our sample to increase sewerage capital spending up until the point where they are in compliance with the CWA's capital mandate, then reduce their own contribution to sewerage capital spending. If this is the case, then the similar relationship between grant size and mandated costs in non-compliant municipalities within and outside of our sample suggests that non-compliant municipalities outside of our sample should have a similar pass-through rate to non-compliant municipalities within our sample. Similarly, since compliant municipalities outside of our sample did not face any mandated costs, we would expect them to have pass-through close to zero.[45]

For the full sample of CWA grants, receiving governments include cities, towns, and sewage districts (Keiser and Shapiro, 2019a), but our sample of municipal spending data only includes cities. Focusing on cities should not affect the generalizability of our results to CWA grants outside of our sample. Towns could use the same crowd-out mechanism as cities. In cases where grants are distributed to special water utility districts that handle wastewater treatment for multiple cities, no single city has authority to divert grant funds received by the utility district to other purposes, and the only diversion possible is for the utility district to pass on grant funds to customers via lower fees, which is consistent with the crowd-out mechanism that we

[44]We test for this more directly in the Appendix. Table 1.15 re-estimates our pass-through results on sub-samples of municipalities divided by tercile of population. While less precise, these estimates are relatively similar to our full sample estimates.

[45]In total, twelve percent of CWA grant dollars were distributed to facilities that were already in compliance with both state and federal treatment technology standards. Assuming that our estimates of the total pass-through rate for compliant and non-compliant municipalities are accurate, we can take a weighted average of pass-through estimates for compliant and non-compliant municipalities to approximate pass-through for all CWA grants. We present the results of this exercise in Table 1.16. This suggests a pass-through rate of 0.353, and we can reject full grant pass-through.

explore in the next section.

1.4 How did Municipalities Spend Crowded-Out Funds?

1.4.1 Redistribution Methods

Evidence of low pass-through in compliant municipalities motivates an examination of how these municipalities adjusted spending in response to grant receipt. Both Callaway and Sant'Anna (2020a) and our stacked difference-in-differences estimator rely on comparisons of newly treated units to not-yet-treated units, but since we observe a relatively small number of compliant municipalities, we leverage a different source of variation to estimate the effect of CWA grants on local spending in compliant municipalities. Since pass-through is generally lower in compliant municipalities than in non-compliant municipalities, we can estimate the effect of grant receipt on spending in compliant municipalities by comparing outcomes between compliant municipalities and non-compliant municipalities before and after grant receipt with equation 1.8.

$$R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it} \tag{1.8}$$

Our dependent variable of interest, R_{it}, is water utility revenue, which we expect to decrease as compliant municipalities redistribute money to residents by lowering utility bills. $grant_{it}$ is a dummy variable that equals one after a municipality receives a grant, $compliant_i$ is a dummy equaling one for compliant municipalities and α_i is a municipality fixed effect. g indexes timing groups, which are defined by the year in which a municipality receives its first CWA grant and α_{gt} is a timing-group-by-year fixed effect.

Including α_{gt} in equation 1.8 lets us estimate the effect of grants to compliant municipalities on spending without using any variation in treatment timing.[46] The result from equation 1.8 is equivalent to estimating the difference-in-difference in equation 1.9 on all observations in timing group g, which compares water revenues between compliant and non-compliant municipalities that receive their first grant in year g in periods before and after g, then averaging together the θ_g^{2X2} from all timing groups.

$$R_{it} = \alpha_0 + \theta_g^{2X2} 1\{t \geq g\} * compliant_i + \delta 1\{t \geq g\} + \omega compliant_i + \varepsilon_{it} \tag{1.9}$$

We first examine the relationship between compliance and water revenue with an event study that com-

[46] As with our grant pass-through results, comparisons of early treated units to late treated units will be unbiased under parallel trends, but comparisons of late treated units to early treated units may be wrong signed. While there is nothing wrong with the first type of comparison, it is likely that variation in compliance will identify a different treatment effect than timing variation, so combining these sources of variation in one equation will reduce the precision of our estimates. Instead, we leverage these two sources of variation with different designs. We present results from a complementary design that identifies off of variation in treatment timing in Appendix Section 1.6.2.1.

pares water revenue in compliant municipalities to water revenue in non-compliant municipalities before and after grant receipt with equation 1.10,

$$R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it} \tag{1.10}$$

then summarize this event study with equation 1.8.[47]

Table 1.8: Non-Compliant Cities are Valid Counterfactuals for Compliant Cities

	(1) Grant Amount	(2) Sewerage Capital	(3) Water Operations Cost
Grant X Compliant	-0.501	-4.786*	-0.513
	(2.967)	(2.450)	(0.907)
Observations	4752	4752	4752

Standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. The dependent variable is grant amount in column 1, sewerage capital spending in column 2, and water operations costs in column 3.

Since we only rely on comparisons between compliant and non-compliant municipalities, the identifying assumption of this design is that, if not for the differences in compliance, average water revenues would have trended similarly in compliant and non-compliant municipalities after grant receipt. Table 1.8 supports comparing spending outcomes between compliant and non-compliant municipalities. Column 1 estimates equation 1.8 with grant amount as the dependent variable. The result is small and insignificant, which shows that compliant and non-compliant municipalities receive similarly sized grants. Column 2 re-estimates equation 1.8 with sewerage capital spending as the dependent variable. Even though grants to compliant and non-compliant municipalities are similar in size, per capita sewerage capital spending is \$4.79 lower in compliant municipalities than in non-compliant municipalities after grant receipt. In column 3, we check if changes in revenue are driven by falling costs in compliant municipalities by estimating equation 1.8 with water operations costs as the dependent variable. The coefficient is small and insignificant, indicating that grants are not associated with changes in the cost of operating water utilities between compliant and non-compliant municipalities

[47] In our crowd-out event studies, we report coefficients for 16 years before and 11 years after grant receipt, which allows us to report only balanced coefficients. These specifications also includes bins for 17 or more years before grant receipt and 12 or more years after grant receipt, but our results are not sensitive to this choice of binning. We examine how effects evolve in later years in Appendix Section 1.6.2.3.

1.4.2 Redistribution Results

Figure 1.7 plots the π_y and γ_y from estimating equation 1.10 with per capita water revenue as the dependent variable. The null estimates in the pre-treatment period provide evidence of parallel trends in water revenues in compliant and non-compliant municipalities prior to grant receipt. The estimates begin to decrease after grant receipt and continue to fall for 11 years after treatment. The gradual decrease in water revenue is consistent with municipalities receiving multiple grants that they could spend over several years.

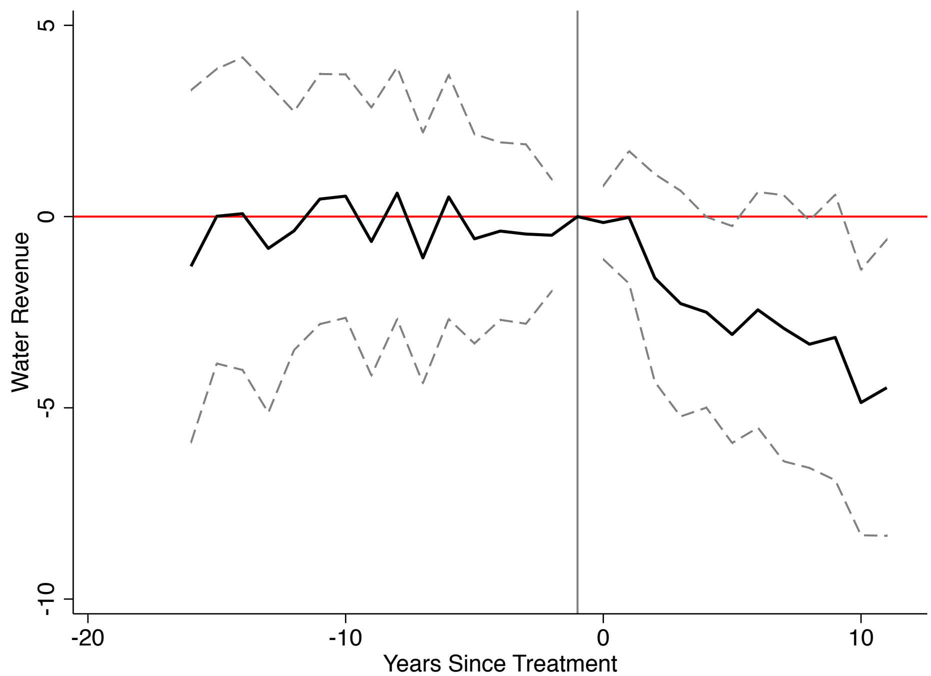


Figure 1.7: Water Revenue Decreased After Grant Receipt in Compliant Municipalities

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ with per water revenue as the dependent variable. $compliant_i$ equals one for observations of compliant municipalities. We include municipality and timing-group-by-year fixed effects, α_i and α_{gt}, where g indexes the year in which a municipality received its first CWA grant.

Column 1 of Table 1.9 summarizes this effect by estimating equation 1.8 with water revenue as the dependent variable. This estimate shows that per capita water revenue decreases by \$2.32 per person in compliant municipalities after grant receipt relative to non-compliant municipalities that receive grants in the same year. Compared to our sample's median per capita water revenue of \$23, this estimate represents a 10 percent decline in water revenues raised by compliant municipalities in response to grant receipt.[48]

Columns 2-4 of Table 1.9 present alternative specifications. The shape of the event study in Figure 1.7

[48] Figure 1.11 checks if this effect is consistent across timing groups by presenting the θ_g^{2X2} from estimating equation 1.9 on each timing group separately. All of the estimates in this figure are negative and relatively close to the result in column 1 of Table 1.9, so no particular within-timing group comparison is driving this result. This figure does not include coefficients for the 1980, 1985 or 1988 treatment cohorts because either no compliant or no non-compliant units were treated in those years. We omit the coefficient for the 1978 cohort because only one compliant and one non-compliant municipality became treated in that year.

suggests that this figure might be better summarized with a trend-break specification, so column 2 estimates a version of equation 1.9 where we interact the treatment dummy with year relative to treatment . This estimate suggests that water revenues decrease by an additional \$0.36 each year after treatment. The map in Figure 1.2 shows that the estimates in columns 1 and 2 of Table 1.9 rely on comparisons across large geographic areas, so in column 3, we re-estimate the specification in column 2 with region-by-year fixed effects.[49] This estimate, identified off of within-region variation in compliance, is not meaningfully different than the estimate in column 2, and our result is still strongly significant despite losing a substantial amount of variation. Finally, column 4 shows that our results are robust to including year fixed effects in place of cohort-by-year fixed effects.[50]

Table 1.9: Compliant Cities Spent Crowded Out Funds on Lowering Water Bills

	(1) Water Revenue	(2) Water Revenue	(3) Water Revenue	(4) Water Revenue
Grant X Compliant	-2.323**			
	(1.157)			
Grant X Compliant X e		-0.364***	-0.396***	-0.441***
		(0.138)	(0.150)	(0.132)
Timing Group X Year FE	X	X	X	
Region X Year FE			X	
Observations	4752	4752	4752	4840

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates the reduced form effect of grants to compliant facilities on water revenues. Column 1 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Column 2 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i * e_{it} + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ where e_{it} indicates year relative to treatment. Column 3 re-estimates the result in column 2 with region-by-year fixed effects. Column 4 re-estimates the result in column 2 with year fixed effects instead of timing-group-by-year fixed effects.

These results show that water revenues decreased in compliant municipalities relative to non-compliant municipalities after grant receipt. This suggests that compliant municipalities redistributed grant money to residents by lowering water utility prices, consistent with public utility providers offsetting water and sewerage utility costs with CWA grant money in order to lower prices to consumers.

1.5 Discussion & Conclusion

There was substantial heterogeneity in the effect of CWA grants on sewerage capital spending. Municipalities spent grant money on capital upgrades up to the point where they were in compliance with the CWA's new treatment technology requirements, but after municipalities met these requirements, or if none were in place,

[49] We divide the United States into four regions; the Northeast, consisting of CT, MA, NH, NJ, NY, PA and RI, the Midwest, consisting of IA, IL, IN, KS, MI, MN, MO, NE, OH and WI, the South, consisting of AR, AL, DE, FL, GA, LA, MD, MS, NC, OK, SC, TN, TX, VA and WV and the West consisting of AZ, CA, CO, MT, NM, OR, UT and WA. States not listed do not appear in our data.

[50] Figures 1.12 and 1.13 present the event studies associated with columns 3 and 4 of Table 1.9.

grants crowded out money that municipalities were already spending on sewerage capital. Per capita water revenues decreased by ten percent from the median after grant receipt in municipalities with low pass-through relative to municipalities with higher pass-through, suggesting that municipalities redistributed crowded out funds to residents through reductions in water bills. In total, receiving governments in our sample spent 44.8% of all CWA grant dollars on sewerage capital.

Given previously quantified benefits of the CWA, accounting for low pass-through can lead to favorable benefit to cost ratios of CWA-funded capital upgrades. Dividing previously estimated benefit/cost ratios of CWA grants by our pass-through estimates suggests that the CWA grant dollars that municipalities spent on sewerage capital have a benefit/cost ratio of 1.01. This result comes from an analysis of $25.5 billion of the $153 billion in grants funded by the CWA. While our sample does not include grants distributed to relatively small municipalities, pass-through is not strongly correlated with population in our sample, which suggests that CWA grants distributed to smaller municipalities have a similar pass-through rate.

Heterogeneity in pass-through informs which types of variation researchers can use to recover the effect of the average grant dollar on spending. Many evaluations of intergovernmental grant programs estimate pass-through using difference-in-differences designs with continuous treatment, but when the effect of a grant of a given size is not constant across receiving governments, estimates identified off of continuous variation in grant size that do not account for heterogeneity in pass-through across units can either over- or under-estimate the effect of the average grant dollar on targeted spending. In the case of the Clean Water Act, estimates of grant pass-through identified off of variation in grant size over-estimate the total pass-through rate.

The relationship between grant amount and targeted spending is likely heterogeneous for many grant programs. As an example, the American Recovery and Reinvestment Act of 2009 distributed almost $50 billion to state departments of education to offset budget shortfalls caused by the Great Recession through the State Fiscal Stabilization Fund. The costs that state governments faced depended on the severity of the local economic downturn, but the federal government determined grant amount based on state population (Superfine, 2011). Budget shortfalls varied widely across states, so if pass-through depended on the costs that receiving governments faced to enact policy, it is unlikely that all states would have responded to transfers of a given size in the same way. More generally, grant size is often positively correlated with receiving governments' preferences for public goods (Knight, 2002), so the effect of a dollar of grant funding on targeted spending is likely increasing in grant size in many settings.

We cannot know for sure if heterogeneity in the relationship between grant amount and targeted spending is the cause of the inconsistent estimates of grant pass-through found in other settings, but future research that accounts for heterogeneity in pass-through across receiving governments (as well as heterogeneity across time relative to treatment in cases where the timing of grant receipt varies) may provide further evidence that this is the case. That being said, researchers are usually limited in the types of variation they can leverage to estimate pass-through; donor governments often distribute grants in such a way that there is no never-treated group and no variation in treatment timing, making it difficult to test for homogeneity in pass-through across receiving governments.

We conclude by noting several implications of this study with respect to the design of intergovernmental grant programs. First, crowd-out may be high in our setting because the CWA allotted grants toward a function of local government that had a direct revenue stream, enabling grant dollars to "drip" from where they hit down to local residents through the low-friction adjustment of water rates. This suggests that crowd-out may be more severe in settings where local governments have a low-friction channel through which to redistribute grants. Key differences between grants that target environmental spending and other intergovernmental grant programs may limit the generalizability of this result; since downstream communities are the primary beneficiaries of environmental investment (e.g. communities downstream from a wastewater treatment plant or communities downwind from air polluters), a city's environmental spending is not salient to constituents, whereas fees (in this case, water bills) are. In many other settings, taxpayers observe their state or local government receiving federal grants for a specific purpose and would notice if targeted spending did not increase, so in most cases both spending and fees are salient to constituents.

Second, pairing grants with new regulation creates a binding constraint, which (weakly) increases the likelihood of full pass-through and the desired policy outcome. Many grant programs use "maintenance of effort" requirements to achieve a similar end, but these requirements are notoriously difficult to enforce. In essence, new regulation can act like a binding maintenance of effort requirement that is relatively easy to enforce, since a donor government can check if a receiving government satisfies a mandate rather than auditing that receiving government's spending. For this method of enacting policy to be effective, donor governments must be able to identify which receiving governments need fiscal assistance and accurately estimate the costs of grant-funded programs for each receiving government.

Similarly, our results suggest that supplementing new regulation with grants can allow local governments to comply with new requirements without reducing spending in other categories (Baicker, 2001; Baicker and Gordon, 2006) or raising additional revenue (Jerch, 2018).

1.6 Appendix

1.6.1 Additional Pass-Through Results

1.6.1.1 Selection on Treated Potential Outcomes

In Section 1.3.4.2, we document heterogeneity in grant pass-through by grant size and compliance with the CWA's treatment technology mandate, but these are not necessarily the only sources of heterogeneity in pass-through. Since states distribute CWA grants to municipalities according to priority lists, we might expect treatment timing to be positively correlated with treated potential outcomes. This is not a threat to identification (we do not place any restrictions on the evolution of outcomes after treatment and our identifying assumption allows for selection into treatment on the basis of unobserved time-invariant characteristics), or the interpretation of the heterogeneity we document in Section 1.3.4.2 (as shown in Figures 1.18 and 1.19, the distribution of treatment timing is similar across compliant and non-compliant municipalities), but it does have implications for the best way to summarize treatment effects. Callaway and Sant'Anna (2020b) suggests that when treatment effects vary across timing groups, we should first aggregate the $ATT(g,t)$ for each timing group g, then combine these timing group-specific effects into a summary measure by taking a weighted average of the effect in each timing group. Specifically, we first calculate

$$\tilde{\theta}(g) = \frac{1}{1985 - g + 1} \sum_{t=1972}^{1987} 1\{g \leq t\} ATT(g,t)$$

for each group, then summarize the effect with

$$\theta_S = \sum_{g=1972}^{1985} \tilde{\theta}(g) P(G = g)$$

θ_S is similar to the simple aggregation based on group-size given by equation 1.2. The difference is that the weights in θ_S do not depend on how long a unit has been treated for, while the weights in equation 1.2 depend on how many periods we observe a timing group for after treatment. This weights up groups that are treated early, so in the presence of selection on gains, equation 1.2 weights up units that experience the largest effects. In contrast, the weights in θ_S only depend on group size.

Table 1.17 presents averages of the $\tilde{\theta}(g)$ for all municipalities in the first three treatment groups, the second three treatment groups and the last five treatment groups in columns 1-3. Column 4 presents θ_S for the entire sample. θ_S is similar to the dynamic and simple aggregations in Panels A and D of Table 1.2, which suggests that treatment timing is not correlated with treated potential outcomes.

Column 5 calculates θ_S with grant amount as the dependant variable. Dividing the result in column 4 by this first stage estimate implies a pass-through rate of 0.46, which is consistent with the results in Table 1.2.

1.6.1.2 Alternative Definitions of Full Pass-Through

Cost-sharing yields two different possible interpretations of pass-through. We define full pass-through as \$1 in grant funding leading to a \$1 increase in sewerage capital spending. Alternatively, since municipalities are required to pay 25% of project costs, one could define full pass-through as \$1 in grant funding leading to a \$1.33 increase in sewerage capital spending (the actual \$1 in grant aid plus the \$0.33 in local obligations). Our estimates in the main text are based on the first definition, appealing to the notion of federal funding itself being spent as intended. We present pass-through estimates using the second definition in Tables 1.18 and 1.19. These tables re-estimate the results in Tables 1.2 and 1.4 after multiplying grant amount by 1.33 in years before 1981, and by 1.82 in 1981 and beyond. Using this definition of full pass-through, we still reject full pass-through for the full sample, non-compliant municipalities that received large grants, and compliant municipalities. We cannot reject full pass-trough or complete crowd-out for non-compliant municipalities that received small grants, but the point estimate for these municipalities suggests that, while CWA grants led to an increase in targeted spending equal to grant amount, they did not cause an increase in local contributions to sewerage capital spending.

1.6.1.3 Average Pass-Through Rate

In the main text, we attempt to recover the total pass-through rate of CWA grants, but that is not the only potential parameter of interest. Figure 1.20 presents the distribution of per capita grant amount for grants in our sample and shows that grant amount is distributed approximately lognormally; given this distribution, it is possible that a few larger grants are meaningfully driving our estimates of the total pass-through rate. Since we want to place the largest weight on the most expensive grants when analyzing the effectiveness of a grant program, this is a desirable feature. That being said, we may instead be interested in the "average pass-through rate", the percent of grant dollars spent as the donor government intended in the average receiving municipality.

We can estimate the average pass-through rate by estimating pass-through in each unit separately, then taking an average across municipalities. To do this, we first estimate equations 1.4 and 1.11 on observations in stack each S.

$$g_{it} = \alpha_0 + \beta_{s_{fs}} post_t * treat_i + \delta post_t + \omega treat_i + \varepsilon_{it} \tag{1.11}$$

We then divide β_s^{2X2} by $\beta_{s_{fs}}$ to obtain an estimate of pass-through for municipality s, then average the resulting within-stack pass-through rates across municipalities.

We present estimates of the average pass-through rate in Table 1.20. While we cannot reject full pass-through or full crowd-out, the point estimate is similar to our estimates of the total pass-through rate.

1.6.1.4 Excluding Municipalities that did not Receive Grants

Since grant timing is determined by a municipality's placement on a state priority list and priority lists are explicitly based on factors that are unlikely to be correlated with trends in sewerage capital spending, we identify our main results off of variation in the *timing* of grant receipt. It is less clear if grant *receipt* is uncorrelated with trends in local spending, but if we are willing to assume that whether or not a municipality received a grant is uncorrelated with trends in local spending, we can use municipalities in our finance data that never receive a CWA grant as an additional control group.

Using never-treated units as a control group only requires minor changes to our estimators. Our Callaway and Sant'Anna (2020a) estimator works in exactly the same way, except we now estimate group-time average treatment effects with

$$\widehat{ATT(g,t)} = E_n[\widehat{C}_t - \widehat{C}_{g-1}|N_i = 1] - E_n[\widehat{C}_t - \widehat{C}_{g-1}|N_i = 0]$$

where N_i is a dummy variable that equals one for never treated units. We can then aggregate these group-time treatment effects together the same way we did in Section 1.3. We present these results in Figures 1.14 and 1.15 and Table 1.21. These estimates are similar to our main results.

For our stacked estimator, we construct stacks so that, instead of using units treated at least seven years after the treated unit as controls, we use never-treated units as controls. We present these results in Table 1.22.

1.6.1.5 Observable Characteristics do not Predict Grant Timing

The identifying variation of our grant pass-through research design comes from *when* each municipality receives its first CWA grant. To identify causal effects with this variation in a difference-in-difference framework, we must assume that the timing of grant receipt is uncorrelated with trends in sewerage capital spending. We perform more explicit tests of our parallel trends assumption in the main text, but in this section, we provide suggestive evidence that early and late treated units would have followed parallel trends in absence of treatment by showing that early and late treated municipalities are observably similar before treatment. We test if observable characteristics of municipalities predict the timing of grant receipt with equation 1.12.

$$GrantYear_i = \alpha_0 + \alpha_1 compliant_i + \alpha_2 X_{ip} + \varepsilon_{ip} \tag{1.12}$$

We regress the year municipality i receives its first CWA grant on observable municipality characteristics

X from a baseline year p, conditional on the municipality not yet receiving a grant in year p. Table 1.23 presents results from estimating equation 1.12 for baseline years 1971, 1974 and 1977.[51] X_{ip} includes city population and per-capita measures of both holistic (e.g. total revenue, total intergovernmental revenue, total capital outlays) and specific (e.g. sewerage capital outlays, water utility revenue) categories of government revenue and expenditure.

None of these observable characteristics consistently predict grant timing. All of the point estimates are small and precisely estimated and the only covariates that have a significant effect on year of grant receipt are water capital spending and log population using 1971 as a baseline year. The coefficients on these variables do not represent economically meaningful changes; a ten percent difference in 1971 population is associated with a municipality receiving its first grant 0.08 years earlier and a one standard deviation increase in water capital outlay is associated with a less than one year change in grant timing. Both of these characteristics are no longer significant and change signs when we use other baseline years. While the covariates are jointly significant when we use baseline years of 1971 and 1974, a standard deviation change in any spending variable in either year represents a less than one year change in grant timing. This shows that grant timing is not strongly correlated with any municipal spending characteristics, which motivates a design that leverages variation in treatment timing to estimate pass-through.

1.6.1.6 Alternative Pass-Through Specifications

In Section 1.3, we summarize pass-through for the full sample using Callaway and Sant'Anna (2020a) and explore heterogeneity in pass-through using a stacked difference-in-differences estimator. We discuss our choice of estimators in greater detail in Appendix 1.6.3, and in this section, we present additional results to show that our results not sensitive to this choice.

Figure 1.16 presents the π_y and γ_y from estimating the following stacked event study on a pooled sample of all stacks.

$$C_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * treat_{is} + \sum_{y=0}^{6} \gamma_y 1\{t - t_i^* = y\} * treat_{is} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its} \tag{1.13}$$

Time relative to treatment is defined by $1\{t - t_i^* = y\}$, which is a dummy variable that equals 1 for observations of municipality i y years before or after t_i^*, the year municipality i receives its first CWA grant and $treat_{is}$ is a dummy variable that equals one for the treated unit in each stack. Similar to the results in Figure 1.4, we see null effects before treatment. The estimates increase after the arrival of a municipality's first CWA grant and remain high for up to 6 years after grant receipt.

[51] Roughly one third of grants to municipalities in our sample were distributed from 1972 to 1973, 1974 to 1976 and post-1976.

Table 1.24 summarizes the results in Figure 1.16. Column 1 estimates the first stage relationship between grant receipt and per-capita grant amount by estimating equation 1.5 on a pooled sample of all stacks. Column 2 estimates the associated reduced form effect of grant receipt on sewerage capital spending by estimating equation 1.6 on all stacks. Column 3 estimates equation 1.6 using D_{cy} as an instrument for grant amount, which returns an estimate equivalent to dividing the reduced form by the first stage. This implies a pass-through rate of 0.342. Using this result, we can reject full pass-through in the full sample (i.e. a test of the hypothesis that the IV estimate in column 3 equals one).

Table 1.25 re-estimates the results in Panel B of Table 1.2 on sub-samples of municipalities. These results are averages of the $\widehat{\theta}_D(e)$ from equation 1.1 for e s.t. $0 \leq e \leq 6$. Overall, these results are similar to those from our stacked difference-in-difference; we can reject full pass-through in non-compliant municipalities that received large grants, but cannot reject full pass-through for non-compliant municipalities receiving small grants. Since the effect of grant receipt on sewerage capital spending is marginally significant for non-compliant municipalities receiving small grants, we can also marginally reject complete crowd-out in these municipalities.

We obtain similar results in Table 1.26, where we re-estimate the results in Table 1.25 using the aggregation given by equation 1.2 (equivalent to Panel D of Table 1.2).

We cannot say anything about compliant municipalities using this estimator. Since there are relatively few compliant municipalities in our sample, the simultaneous critical value is arguably 'too large' to be reliable when we estimate the effect of grant receipt on sewerage capital spending in these municipalities, so we default to our stacked results.

1.6.1.7 Measuring the Benefits of CWA Grants

In this section, we discuss several potential issues with using previously estimated benefits of CWA grants to find the benefit-to-cost ratio of CWA grants that municipalities spent on sewerage capital.

First incorporating other currently unmeasured benefits could further increase the benefit-to-cost ratio CWA grants, and a full welfare analysis would include these unmeasured benefits.

Second, a more formal analysis would take the confidence intervals and potential bias of estimates of the benefits of CWA grants into account. In particular, both the hedonic estimates in Keiser and Shapiro (2019a) and the infant health estimates in Flynn and Marcus (2021) come from TWFE estimators with continuous treatment, but we show in Section 1.3.5 that this type of estimator can produce misleading results in the presence of heterogeneity in treatment effects across compliance or heterogeneity in treatment effects across time. The main estimates in Flynn and Marcus (2021) account for heterogeneity across compliance. Flynn and Marcus (2021) also obtains similar results when using estimators that account for dynamic treatment effects. While it does not explicitly account for heterogeneity across time, the event study graphs in Keiser and Shapiro (2019a) do not show evidence of dynamic treatment effects, and instead show that CWA grant receipt led to a mean shift in housing prices. Keiser and Shapiro (2019a) does not account for heterogeneity across compliance either, but since there is evidence that similar demographic groups sorted into communities downstream from compliant and non-compliant municipalities (see Table 2 in Flynn and Marcus (2021)), it is not clear that the effect of CWA grants on housing prices is heterogeneous across compliance. That being said, if this heterogeneity is present, it could bias estimates of the effect of CWA grants on housing prices. It is not clear a priori which direction this bias would move in.

Finally, the benefits of CWA grants may be driven in part by spending on categories besides sewerage capital. In Section 1.4, we provide evidence that compliant municipalities redistributed grant dollars to residents through reduced water bills. A hedonic model suggests that these savings should be capitalized into housing prices. Since Keiser and Shapiro (2019a) estimates the effect of CWA grant receipt on housing prices by comparing all housing units within 25 miles of treated plants against all housing units within 25 miles of untreated plants, the estimated benefit-to-cost ratio of 0.26 in Keiser and Shapiro (2019a) may be driven in part by the value that residents place on reduced water bills. Consequently, our estimated benefits likely include the value that residents place on lowered water rates in addition to the benefits of improved infant health and the value that residents place on increased water quality.

Reductions in water rates are driven by grant dollars that municipalities redistributed to residents. For this reason, we do not want to include these benefits when we are trying to find the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital, but we cannot disentangle the portion of the effect of CWA grants on housing prices driven by lower utility prices and the portion driven by improved water quality. We can instead find the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital *and* CWA grant dollars that municipalities redistributed to residents through lowered water bills. Table 1.32 suggests that 48.5% of grant dollars for compliant facilities went to reducing water bills (though this estimate is imprecise). Twelve percent of CWA grant dollars went to compliant municipalities, so this represents 5.82% of all CWA grant dollars. Since the benefit-to-cost ratio of 0.45 from Keiser and Shapiro (2019a) and Flynn and Marcus (2021) comes from both the 44.6% of CWA grant dollars that municipalities spent on sewerage capital and the 5.82% of all CWA grant dollars went to reducing water bills, we can obtain the benefit-to-cost ratio of CWA grant dollars that were spent on sewerage capital or redistributed to residents through reduced water bills by dividing previously estimated benefit-to-cost ratios of CWA grants, 0.45, by the portion of CWA grant dollars spent on sewerage capital and reducing water bills combined, 0.504. This suggests a benefit-to-cost ratio of 0.89, however this is not exactly the number that we are interested in.

We want the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital. We can use the results in Table 1.32 to estimate a range for this number. To generate a lower bound, we can assume that reductions in water rates were completely capitalized into housing prices (i.e. a one dollar reduction in water bills leads to a one dollar increase in housing prices). If this were the case, then, of the $0.26 of the average dollar of CWA grant aid capitalized into housing prices, 0.058 would come from reduced water bills and the other 0.202 would come from improvements in water quality. Combining this number with the infant health benefits from Flynn and Marcus (2021) suggests that the CWA grant dollars spent on everything except reducing water bills would have a benefit-to-cost ratio of 0.392. We can divide this number by the pass-through rate to calculate the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital. This suggests a benefit-to-cost ratio of about 0.879. If instead reductions in water bills were not capitalized into housing prices at all, all of the benefits of the CWA in Keiser and Shapiro (2019a) and Flynn and Marcus (2021) would come from sewerage capital spending, and the benefit-to-cost ratio of CWA grant dollars spent on sewerage capital spending would be $\frac{0.45}{0.446}$, or about 1.01. Where in this range the actual benefit-to-cost ratio of CWA grant dollars spent on sewerage capital spending falls depends on the capitalization rate of reductions in water bills, but the existing literature has not reached a consensus on the effect of changes in fees on housing prices (Yinger et al., 2016). Recent results suggest that there should be near full capitalization if the housing supply is inelastic (Lutz, 2015), but determining the elasticity of housing around municipal wastewater treatment facilities is outside the scope of this paper.

Similarly, the portion of CWA grant dollars that municipalities spent on categories besides sewerage capital spending and lowering water bills may have caused increases in housing prices or improvements in infant health, which would similarly decrease the benefit-to-cost ratio of CWA grants spent on sewerage capital.

1.6.1.8 Baseline Sewerage Spending by Compliance and Grant Size

If municipalities that received small grants had relatively low pre-treatment sewerage capital spending, it could be the case that a relatively small portion of the grants distributed to these municipalities are inframarginal. If less of these grants dollars are inframarginal, then the higher pas-through that we document in municipalities that receive small grants could be mechanical; municipalities that received small grants would not have had much spending to to crowd-out, while municipalities that received larger grants spend more in sewerage capital before grants, and would have had more baseline spending to crowd out. Table 1.27 shows that this is not the case, and that pre-treatment spending is highest in non-compliant municipalities that receive small grants.

1.6.2 Additional Water Revenue Results

1.6.2.1 Estimating Water Revenue Results Using Timing Variation

In Section 1.4, we leverage variation in pre-CWA compliance with the CWA's treatment technology mandate to examine how compliant municipalities adjust spending in response to grant receipt, but we obtain similar results when we estimate these effects using variation in treatment timing.

In Figures 1.21 and 1.22, we estimate separate event studies for compliant and non-compliant municipalities. These event studies compare water revenue after treatment to water revenue in not-yet-treated municipalities using Callaway and Sant'Anna (2020a). Sub-Figure 1.21 shows that grant receipt has no significant effect on water revenues in non-compliant municipalities. Sub-Figure 1.22 shows that water revenues are stable before grant receipt and decrease after grant receipt in compliant municipalities. Table 1.28 summarizes these figures.

Note that the estimates in column 1 of Table 1.28, which are identified off of comparisons between early and late treated compliant municipalities, are larger than the estimates in Table 1.9, which are identified off of comparisons between compliant and non-compliant municipalities. Similarly, the event study in Figure 1.22 has a somewhat different shape than the event study in Figure 1.7. This does not reflect a problem with either design. Rather, it demonstrates that these two sources of variation identify different treatment effects.

This difference is caused in part by the way in which OLS combines the effects in different treatment cohorts into summary measures. OLS weights based on the size of timing groups and within-group variance of the treatment dummy, which artificially weights up units treated near the middle of the panel (Goodman-Bacon, 2021a). To avoid weighting based on within-group variance, we can take a weighted average of the effect in each timing group where the weights are based only on the size of each timing group (i.e. a weighted average of the effects in Figure 1.11). This average is -3.76, which is closer to the estimate in column 1 of Table 1.28.

1.6.2.2 Why was there No Effect in Non-Compliant Municipalities?

Table 1.28 shows that grant receipt had no average effect on water revenue in non-compliant municipalities. This result provides support for the identifying assumption of the design we use in Section 1.4, but it is somewhat counter-intuitive; since we do not estimate full pass-through in all non-compliant municipalities in our sample, we might expect grants to cause a decrease in water revenues in non-compliant municipalities.

Figure 1.1 suggests one explanation for this result. Note that while, on average, small grants are similar to costs and large grants are greater than costs, there are non-compliant municipalities that receive grants that are too small or too large along the entire distribution of grant size. If this leads some municipalities redistribute grant funds by decreasing their water bills while other municipalities increase water bills to cover costs greater than grant amount, these effects will cancel out.

Figures 1.23 and 1.24 provides suggestive evidence that this is the case by re-estimating the results in Figure 1.21 on sub-samples of municipalities where grants are larger or smaller than estimated costs. Figure 1.23 suggests that municipalities that receive grants that are smaller than costs increase their water bills in order to raise the revenue needed to come into compliance with the CWAs treatment technology mandate. Figure 1.24 does not provide clear evidence of an increase or decrease in water revenues in non-compliant municipalities that receive grants larger than estimated costs.

Table 1.29 summarizes the results in Figures 1.23 and 1.24. The estimate for municipalities that receive grants smaller than costs is positive, while the estimate for municipalities that receive grants larger than costs is negative, as expected, though neither of these effects are statistically significant. This suggests that grants did not cause a decrease in water revenues in non-compliant municipalities because these countervailing effects cancel out.

The results in Table 1.28 also motivate a discussion of how our results relate to Jerch (2018). Jerch (2018) uses an instrumental variables approach to estimate the effect of the CWA's capital mandate on local spending and finds that the capital mandate imposed a large financial burden on non-compliant municipalities, and that non-compliant municipalities then increased water bills in response to the capital mandate. Two important differences between this paper and Jerch (2018) can explain why we obtain different results when we estimate the effect of grant receipt on water revenue for non-compliant municipalities: first, the research design in Jerch (2018) intentionally does not capture the effect of CWA grants, and second, the sample of municipalities that Jerch (2018) uses includes municipalities that did not receive any CWA grants. In contrast, we only look at the effect of the CWA's capital mandate on water revenues in municipalities that received CWA grants, and find that, on average, non-compliant municipalities that received CWA grants did not change their water bills. This suggests that the increase in water revenues in non-compliant municipalities documented in Jerch (2018) are driven by municipalities that did not receive CWA grants. This is consistent with other studies of unfunded federal mandates (Baicker, 2001; Baicker and Gordon, 2006).

1.6.2.3 How Long Do Effects Persist?

Figures 1.7 and1.22 both show that the effect of grants to compliant municipalities on water revenue grow over time, which motivates an examination of how long these effects persist. Our spending data includes observations from 1956 to 1999 and the latest treated municipality receives its first CWA grant in 1988. This allows us to estimate event study coefficients for 12 balanced post-treatment periods in our main specifications, but if we drop the compliant municipalities treated after 1981, we can estimate event study coefficients for 19 balanced post-treatment periods.

Figure 1.25 re-estimates the results in Figure 1.7 on this sample. The estimates for the 16 years before treatment and the first 12 years after treatment are very similar to those in Figure 1.7. The effect of grant receipt on water revenue reaches its lowest point 13 years after treatment, then begins to return to pre-treatment levels.

This pattern of effects is consistent with municipalities receiving multiple grants that were fungible across years. Figure 1.26 shows the distribution of grants that a municipality received each year relative to the year the municipality receives its first grant. This figure shows that some municipalities receive CWA grants up to 16 years after initial grant receipt, which can explain why the effect of grant receipt on water revenues persists well into the post-treatment period.

1.6.2.4 Alternative Water Revenue Data

Starting in 1967, the Census began collecting spending data from a larger sample of municipalities once every five years. Since it only contains one pre-treatment period, we do not use this data in our main crowd-out specifications, however, we obtain similar results when we re-estimate the results in Sections 1.4 and 1.6.2.1 using this dataset.

Figure 1.27 re-estimates the results in Figure 1.7 on this data. The pattern of effects is similar to those in Figures 1.7 and 1.25. Water revenue decreases in compliant municipalities relative to non-compliant municipalities for 15 years after grant receipt, then begin to return to pre-treatment levels.

Table 1.30 summarizes Figure 1.27. Column 1 estimates equation 1.8 on our alternative dataset. Column 2 re-estimates the specification in column 1 with region-by-year fixed effects, and column 3 re-estimates the specification in column 1 with a year fixed effect in place of a timing-group-by-year fixed effect. The results are similar to the estimate of equation 1.8 in column 1 of Table 1.9.

Figures 1.28 and 1.29 and Table 1.31 then re-estimate the results in Figures 1.21 and 1.22 and Table 1.28 respectively. The results are again similar to those we obtain with our main dataset.

1.6.2.5 How Much Crowd-Out can Redistribution Account For?

In Section 1.4, we show that, relative to non-compliant municipalities that received grants in the same year, compliant municipalities reduced water bills by \$2.32 per capita after grant receipt. In this section, we examine how much crowded-out spending reductions in water revenue account for.

Column 1 of Table 1.32 presents estimates of equation 1.8 with sewerage capital spending as the dependent variable. Panel A estimates this equation with a dummy treatment variable and Panel B re-estimates this specification interacting the treatment dummy with year relative to treatment. These results suggest that, after grant receipt, compliant municipalities increased sewerage capital spending by \$4.76 per capita (or by an additional \$0.51 per year) less than non-compliant municipalities that were treated in the same year. Since, as shown in column 1 of Table 1.8, compliant and non-compliant municipalities that were treated in the same year received grants of similar size, this difference in sewerage capital spending is likely due to differences in pass-through. For this reason, dividing our reduced form estimates of the effect of grant receipt on water revenues in compliant municipalities from Table 1.9 (also presented in column 3 of Table 1.32) will show how much crowded-out spending reductions in water revenue can account for. We present these results in column 3 of Table 1.32. While they are imprecise and we cannot reject that the coefficients equal 1 (which would suggest that reductions in water revenue can account for all crowded-out spending) or 0 (which would suggest that they do not account for any crowded out spending), taken at face value, these results suggest that reductions in water revenue can account for between half and three-quarters of crowded-out spending.

Since we do not find evidence of an increase in targeted spending in compliant municipalities, any spending that this does not account for was likely redistributed to other functions of local government according to local preferences.

1.6.3 Choice of Estimators

There are many ways to estimate treatment effects in a difference-in-differences framework, including two way fixed effects estimators with binary and continuous treatment, Callaway and Sant'Anna (2020a), and stacked difference-in-difference estimators. Each of these methods identifies off of different comparisons and combines these comparisons into summary measures according to different weights. The decision of which method to use is context specific, so we rely on different estimators for different portions of our analysis. In this section, we discuss our choice of estimators in detail.

1.6.3.1 Two Way Fixed Effects with Binary Treatment

We could leverage variation in grant timing with a two way fixed effects (TWFE) estimator. To do this, we would first estimate the reduced form effect of grant receipt on sewerage capital spending with the following equation

$$C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{1.14}$$

where C_{it} is per capita sewerage capital spending, D_{it} is a dummy variable that equals one for observations of municipalities after they receive their first grant and α_i and α_t are municipality and year fixed effects. We would then divide our estimate of β_{rf}^{TWFE} by an estimate of the first stage relationship between grant receipt and per capita grant amount, g_{it}, from estimating

$$g_{it} = \alpha_0 + \beta_{fs}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{1.15}$$

Since every unit in our sample is eventually treated, estimating β_{rf}^{TWFE} returns an average of comparisons between (1) newly treated municipalities relative to municipalities that have not yet been treated and (2) newly treated municipalities relative to already-treated municipalities. When treatment effects are dynamic, municipalities still *actively responding to treatment* are not a valid counterfactual to represent potential outcomes in the *absence of treatment*, so the second type of comparison can produce results that do not have a causal interpretation. Mechanically, changes in the treatment effects of already-treated units over time are subtracted from a TWFE estimate and we cannot interpret the resulting coefficient as the average treatment effect of the treated (ATT) (Goodman-Bacon, 2021a).

We have an a priori reason to expect the effect of CWA grants on sewerage capital spending to change over time. States distributed CWA grants to fund specific projects, so we would expect grant receipt to cause an increase in sewerage capital spending that only lasts until the project is completed. When spending returns to pre-treatment levels, this decrease is subtracted from a TWFE estimate, biasing the estimate upwards.

We illustrate this problem with a simple example (shown graphically in Figure 1.30). Consider a setting where there are two municipalities, A and B, and three time periods, $t = \{0,1,2\}$. Treatment turns on for municipality A in period $t = 1$ and for municipality B in period $t = 2$. After treatment turns on, the outcome variable y increases by one, then returns to pre-treatment levels in the next period. Summarizing this effect with TWFE will return an average of (1) a comparison of municipality A to municipality B in periods 0 and 1 where municipality A is the treated unit and (2) a comparison of municipality B to municipality A in periods 1 and 2 where municipality B is the treated unit. If we denote y_i^t as the outcome variable in municipality i in period t, the first comparison will be

$$(y_A^1 - y_A^0) - (y_B^1 - y_B^0) = (4-3) - (1-1) = 1$$

and the second will be

$$(y_B^2 - y_B^1) - (y_A^2 - y_A^1) = (2-1) - (3-4) = 2$$

TWFE will return a weighted average of these two comparisons, which is strictly greater than the true effect, and our result will be biased upwards. This sort of upward bias could be one reason that estimates of grant pass-through are higher than economic theory would predict in situations where grants roll out over time.

To test the validity of a TWFE estimator in this context, we check how much of our estimate of β_{rf}^{TWFE} comes from comparisons of newly treated units relative to already-treated units. We first estimate equation 1.14 on the full sample, then decompose the resulting estimate of β_{rf}^{TWFE} into every possible comparison between timing groups using Goodman-Bacon et al. (2019). Figure 1.32 presents a scatterplot of these comparisons and their associated weights. Table 1.33 summarizes Figure 1.32 and shows that more than half of our TWFE estimate comes from comparisons of newly treated units to already-treated units, raising concerns over its interpretation and motivating a different estimator.

1.6.3.2 Dose-Response Two Way Fixed Effects

Researchers often estimate pass-through by leveraging continuous variation in grant size. Callaway et al. (2021) documents several potential problems that arise when using this type of variation in a difference-in-differences framework. In this section, we discuss how the results from Callaway et al. (2021) apply to estimating grant pass-through.

Consider the following equation, which regresses targeted spending on a continuous measure of grant size and unit and time fixed effects:

$$C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it} \tag{1.16}$$

When there is no never treated group, then, similar to a TWFE estimate with binary treatment, estimating β_{dr}^{TWFE} returns a weighted average of (1) comparisons of newly treated municipalities relative to municipalities that have not yet been treated, (2) comparisons of newly treated municipalities relative to already-treated municipalities, *and* (3) comparisons of municipalities that received larger grants relative to municipalities that received smaller grants. As in Section 1.6.3.1, the first type of comparison produces meaningful estimates under parallel trends. In the presence of dynamic treatment effects, the second type of comparison can produce results that do not have a causal interpretation.

To identify meaningful effects with the third type of comparison, we must assume that the effect of a grant of a given size is homogeneous across receiving municipalities. To see why, consider two groups of municipalities treated in the same period. Municipalities in one group receive grants of size d and municipalities in the the other group receive grants of size d'. Express the effect of d on the group that received grants of size d as $ATT(d|d)$. Under parallel trends, comparing the effects of grants in these groups of municipalities yields the following result:

$$ATT(d|d) - ATT(d'|d') = \underbrace{ATT(d|d) - ATT(d'|d)}_{\text{effect of going from d' to d}} + \underbrace{ATT(d'|d) - ATT(d'|d')}_{\text{selection bias}}$$

We can interpret the first term as the effect of increasing grant size from d' to d on units that actually received d, which is a causal parameter. The problem is that this type of comparison also includes a selection bias term, which will be averaged into our estimate of β_{dr}^{TWFE}.

Under parallel trends, we can write the selection term as

$$E[Y_t(d') - Y_{t-1}(0)|D = d] - E[Y_t(d') - Y_{t-1}(0)|D = d']$$

which is equal to the effect of a grant of size d' on units that actually received d minus the effect of a grant of size d' on units that actually received d'. If we assume that the effect of d' is homogeneous, then these terms cancel.

Since grants of a given size often have different effects in compliant and non-compliant municipalities, this assumption is unlikely to hold in the context of CWA grants. The following example illustrates why heterogeneity in grant pass-through by compliance presents a problem for estimation. Consider what happens when we compare a non-compliant municipality that received a grant of \$210 and increased spending by \$75 to a compliant municipality that received a grant of \$200 and increased spending by \$25. The total pass-through rate of the grants to these two municipalities is 0.244, but comparing these two municipalities implies that each dollar of grant money led to a $\frac{75-25}{210-200} = 5$ dollar increase in targeted expenditure. Comparisons of compliant municipalities to non-compliant municipalities will face the opposite problem; this type of comparison suggests that grants led to a *decrease* in spending, which will bias β_{dr}^{TWFE} downward.

We can address problems caused by heterogeneity across compliant and non-compliant municipalities by estimating β_{dr}^{TWFE} on each of these groups separately, (we present these results in Table 1.34) however, these results may not return the total pass-through rate if treatment effects are dynamic. We illustrate how comparisons across timing groups complicate estimation with a modified version of the example in Figure 1.30 (shown graphically in Figure 1.31). There are two municipalities, A and B, and three time periods, t = $\{0,1,2\}$. Treatment turns on for municipality A in period $t = 1$ and for municipality B in period $t = 2$. Municipality A receives a dose of size 2 and municipality B receives a dose of size 1. After treatment turns on, the outcome variable y increases by dose size before returning to pre-treatment levels in the next period. A TWFE estimate will include (1) a comparison of municipality A to municipality B in periods 0 and 1 where municipality A is the treated unit, divided by the dose size in municipality A, (2) a comparison of municipality B to municipality A in periods 1 and 2 where municipality B is the treated unit, divided by the dose size in municipality B and (3) a comparison of municipality A to municipality B in periods 0 and 3 where municipality A is the treated unit, divided by the the difference in doses between municipality A and municipality B. If we denote the dose size in municipality i as d_i, the first comparison will be

$$\frac{(y_A^1 - y_A^0) - (y_B^1 - y_B^0)}{d_A} = \frac{(5-3)-(1-1)}{2} = 1$$

the second will be

$$\frac{(y_B^2 - y_B^1) - (y_A^2 - y_A^1)}{d_B} = \frac{(2-1)-(3-5)}{1} = 3$$

and the third will be

$$\frac{(y_A^2 - y_A^0) - (y_B^2 - y_B^0)}{d_A - d_B} = \frac{(3-3)-(2-1)}{2-1} = -1$$

Only the comparison of early treated units to not-yet-treated units accurately describe the relationship between dose size and the outcome variable. Note that dose-response variation returns a result that is wrong signed *even though the effect of grants is homogeneous in grant size*.[52]

To interpret results identified off of dose-response variation as the ATT when treatment timing varies, we must assume that the response to a given does in a given time period is constant across groups (Callaway et al., 2021). Since treatment effects are heterogeneous across both time relative to treatment (as shown in Figure 1.4) and compliance (as shown in Figures 1.5 and 1.6), this assumption does not hold. For this reason, our research design does not leverage variation in grant size.

1.6.3.3 Callaway and Sant'Anna

Since it focuses on combinations of $ATT(g,t)$, Callaway and Sant'Anna (2020a) is far simpler than a TWFE estimator. In practice, $\widehat{ATT(g,t)}$, the sample analogue of $ATT(g,t)$, is equivalent to the OLS estimate from a two group, two time period difference-in-difference that compares observations from group g and observations from groups not yet treated at time t in periods t and $g-1$. That is, $\widehat{ATT(g,t)}$ is identical to an estimate of β from equation 1.17.

$$C_{it} = \alpha_0 + \alpha_1 G_g + \alpha_2 1\{T=t\} + \beta(G_g * 1\{T=t\}) + e \tag{1.17}$$

[52]See Theorem 5 in Callaway et al. (2021) for a formal discussion.

Our 12 timing groups, observed across many post-treatment periods, yield 96 different $\widehat{ATT(g,t)}$. Most of the $\widehat{ATT(g,t)}$ are identified off of relatively few observations, so instead of interpreting individual $\widehat{ATT(g,t)}$, we summarize the effect of treatment by aggregating the $\widehat{ATT(g,t)}$ together. Unlike a TWFE estimator, where comparisons are always weighted together based on group size and within-group variance of the treatment dummy (Goodman-Bacon, 2021a), researchers can select weights appropriate to the setting. We report the dynamic aggregation given by equation 1.1 and the simple aggregation given by equation 1.2 in the main text, as well as an aggregation that allows for selective treatment timing in Appendix Section 1.6.1.1.

Callaway and Sant'Anna (2020a) constructs standard errors with a multiplier bootstrap procedure.[53] Instead of re-sampling observations, each bootstrap draw perturbs the influence function of the estimate (which measures the dependence of the estimate on each cluster in the sample). We report multiplier bootstrap standard errors for our summary measures of the first stage relationship between grant receipt and grant amount and the reduced form effect of grant receipt on sewerage capital expenditure.

Rather than derive the influence function for our pass-through estimator, we construct standard errors for our pass-through estimates with a pair-bootstrap procedure. To do this, we take a random sample (with replacement) of municipalities and jointly estimate our first stage and reduced form estimates. We then divide our reduced form estimate by the associated first stage estimate to obtain β_b^{boot}, where b indexes bootstrap iterations. We repeat this process 1000 times and use the results to calculate bootstrap standard errors with the following equation

$$\sqrt{\frac{1}{999}\sum_{b=1}^{1000}(\beta_b^{boot} - \beta^{boot})^2} \tag{1.18}$$

where $\beta^{boot} = \frac{1}{1000}\sum_{b=1}^{1000}\beta_b^{boot}$.

Because pair-bootstrap draws do not preserve the relative size of each timing group, the standard errors for our pass-through estimates will not be efficient if the effect of treatment varies across treatment groups. The results in Section 1.6.1.1 suggest that this is not the case, so changes in the size of timing groups across bootstrap samples should not add variance to our pair-bootstrap estimates.[54]

1.6.3.4 Stacked Difference-in-Differences

When examining heterogeneity in the relationship between grant amount and targeted spending, we rely on estimates from a stacked difference-in-differences estimator. While most of our results are robust to using Callaway and Sant'Anna (2020a) instead, a stacked estimator allows us to define our treatment and con-

[53] See Algorithm 1 in Callaway and Sant'Anna (2020b).

[54] Note that we do not go so far as to assume that treatment timing is truly random. When that is the case, difference-in-differences designs do not provide the most efficient estimates (Roth and Sant'Anna, 2021).

trol groups more flexibly. While Callaway and Sant'Anna (2020a) estimates the average effect of grant receipt on *all units* in a *given timing group* in *different periods*, we construct our stacks so that we estimate the average effect of grant receipt on *individual units* across *all post treatment periods*. This allows us to semi-parametrically examine the relationship between grant amount and sewerage capital spending before constructing summary measures of pass-through. Since timing groups contain both compliant and non-compliant municipalities and municipalities that received different sized grants, we cannot repeat this process with individual $ATT(g,t)$ without making a priori assumptions about the relationship between grant amount and sewerage capital spending, as well as the relationship between compliance and sewerage capital spending. That being said, Callaway and Sant'Anna (2020a) is much more careful about aggregating these results together and conducting inference,[55] which is why we rely on Callaway and Sant'Anna (2020a) for our main results.

A stacked estimator can be more complicated than Callaway and Sant'Anna (2020a), so in this section, we discuss how we construct stacks, as well as inference and identification with a stacked estimator.

Each stack S is defined by a single municipality, called municipality s, which is labeled as treated in stack S. Stack S also includes municipalities that receive grants at least seven years after the treated municipality, which are labeled as controls. Municipalities that receive grants early (specifically, those treated before 1979) are only in stacks in which they are the treated unit. Municipalities that receive grants too late to have any controls of their own (those treated after 1981) are only included as controls in other stacks. Municipalities that receive grants in the middle of our study period (from 1979 to 1981) can be a treated unit in one stack, while observations of that municipality from before it received a CWA grant can be controls in other stacks.

We can estimate the reduced form effect of grant receipt on sewerage capital spending for municipality s, denoted β_s^{2X2}, by comparing spending in municipality s to spending in the control municipalities in stack S before and after t_s^*, the year in which municipality s becomes treated. For a concrete example, consider a municipality treated in 1972. In the stack defined by that municipality, we can estimate a simple two group difference-in-difference where treatment turns on for the treated group in 1972. This compares sewerage capital spending in the treated municipality to spending in the control municipalities before and after 1972. The control group consists of every municipality treated in 1979 or later. This stack will include observations of both the treated and control units in each year from 1956 to 1978. Since there are no observations from 1979 or later in this stack, treatment never turns on for the control group.

Using units treated at least seven years after municipality s as controls is not an arbitrary choice. As shown in Figures 1.3 and 1.4, both grant amount and sewerage capital spending return to near pre-treatment

[55] Callaway and Sant'Anna (2020a) is also more careful about controlling for confounders, though our analysis does not have any controls.

levels by 6 years after treatment. If we only used units treated more than seven years in the future as controls, we could look at a longer post period, but our control group would be very small. We could have a larger control group if we used a shorter post period, but the resulting estimates would not capture increases in sewerage capital spending caused by CWA grants that took several years to appear in the data. Using units treated seven years in the future as controls is a compromise between these two options that likely captures the majority of the effects of CWA grants.

Our estimates of the β_s^{2X2} are noisy, so instead of interpreting individual β_s^{2X2}, we create summary measures of the effect of grant receipt on sewerage capital spending by aggregating the β_s^{2X2} together. Before doing so, we examine heterogeneity in pass-through by both grant size and compliance in Figures 1.5 and 1.6. On the bias/variance trade-off, these figures are very low bias, and we can use them to determine the best way to obtain lower variance summary measures of pass-through. The shapes of these figures motivate us to separately summarize pass-through for compliant municipalities, non-compliant municipalities that received small grants and non-compliant municipalities that received large grants.

We summarize the effect of grants on sewerage capital spending for each of these groups with $\beta_{rf}^{stacked}$ from equation 1.6. In practice, $\beta_{rf}^{stacked}$ is a weighted average of the β_s^{2X2} from each stack. The weights on each β_s^{2X2} in our estimate of $\beta_{rf}^{stacked}$ are based on stack size and within-stack variance of D_{it} (Goodman-Bacon, 2021a). We construct our stacks in such a way that each β_s^{2X2} is weighted equally by our regression. To do this, we collapse our control observations in each stack to yearly means. In the collapsed data, each stack contains two observations for each year (one treated observation and one control observation) for 16 years of pre-treatment data and 7 years of post-treatment data. Since each stack has the same number of observations and the treated unit in each stack is treated for the same percent of the time, each stack has the same weight.

We cannot construct standard errors clustered at the municipality level using this collapsed data. Instead, we construct pair-bootstrap standard errors. To do this, we take a random sample (with replacement) of observations in the unstacked data, clustering at the municipality level to preserve any dependence of error terms across time. We then re-form our stacks and perform our estimates. We repeat this process 1000 times and use the results to calculate bootstrap standard errors with equation 1.18.

In the main text, we assume that treatment timing is uncorrelated with trends in local spending, but we can state the assumption needed to interpret our stacked estimates as the ATT more precisely. Remember that the coefficient from our stacked estimator is an unweighted average of the difference in sewerage capital spending between treated and control municipalities in each stack before and after treatment. Estimating this simple difference-in-difference in a given stack S (defined by municipality s) yields β_s^{2X2}, which we can express in terms of means as

$$\beta_s^{2X2} = (\bar{y}_s^{post} - \bar{y}_s^{pre}) - (\bar{y}_{\sim s}^{post} - \bar{y}_{\sim s}^{pre})$$

This is the difference in the outcome variable in municipality s before and after t_s^* minus the difference in the outcome variable in the control municipalities in stack S before and after t_s^*. If we define $y_{it}(1)$ as the potential outcome if municipality i is treated in year t, and $y_{it}(0)$ as the potential outcome if municipality i is not treated in year t, then we can express β_s^{2X2} in terms of potential outcomes as

$$\begin{aligned}
\beta_s^{2X2} =&(E[y_{it}(1)|c = s, post] - E[y_{it}(0)|i = s, pre]) \\
&- (E[y_{it}(0)|i \neq s, post] - E[y_{it}(1)|i \neq s, pre]) \\
=&E[y_{it}(1) - y_{it}(0)|i = s, post] + E[\Delta y_i(0)|i = s] - E[\Delta y_i(0)|i \neq s] \\
=&ATT_s + E[\Delta y_i(0)|i = s] - E[\Delta y_i(0)|i \neq s]
\end{aligned}$$

ATT_s is the effect of treatment in municipality s, so β_s^{2X2} consists of a causal parameter and an identifying assumption. In this case, to interpret β_s^{2X2} as the effect of treatment in municipality s, we must assume that, on average, municipality s would have followed parallel trends with the control municipalities in stack S in absence of treatment.

Since $\beta^{stacked}$ (the result from a stacked difference in difference) is an unweighted average of the β_s^{2X2} from each stack, we can express $\beta^{stacked}$ as

$$\begin{aligned}
\beta^{stacked} &= \frac{1}{S}\sum_{s=1}^{S}(ATT_s + E[\Delta y_i(0)|i = s] - E[\Delta y_i(0)|i \neq s]) \\
&= \frac{1}{S}\sum_{s=1}^{S} ATT_s + \frac{1}{S}\sum_{s=1}^{S}(E[\Delta y_i(0)|i = s] - E[\Delta y_i(0)|i \neq s])
\end{aligned}$$

so our identifying assumption is that $\sum_{s=1}^{S}(E[\Delta y_i(0)|i = s] - E[\Delta y_i(0)|i \neq s]) = 0$. Note that this does not require us to assume that parallel trends holds in *every* stack, only that any violations of parallel trends cancel out when we take an average of the effect in each stack.

1.6.4 Data Details

1.6.4.1 Grant Data

We begin with grant data from the EPA's Grant Information Control System, which we obtained through a Freedom of Information Act request. This data contains information on the year that the EPA distributed each grant, which municipality received the grant, the specific wastewater treatment facility the grant was designated for and the amount distributed. Keiser and Shapiro (2019a) uses the same data, and Appendix

Section B.4 of Keiser and Shapiro (2019a) demonstrates its accuracy.

Some grants are explicitly for construction (and include a plant code), while others do not have a plant code. It is unclear to what extent these grants were precisely for upgrading wastewater treatment plants, so we drop grants that did not have a specific facility code. This restricts our sample to 33,429 grants. We also drop grant records that are missing the year in which they were distributed, which drops another 3533 observations, as well as the 475 grants distributed to municipalities outside the contiguous U.S. Some municipalities received grants in 1960-72 through predecessor programs similar to the CWA. At least some of these precursor grants are included in our data from the EPA, but we did not assess the quality or completeness of information of grants from before 1972, so we exclude the 5248 observations of precursor grants from our analysis. This leaves us with records of 25,997 grants to 12,291 facilities.

If municipalities in our sample received any of the grants that we drop, our pass-through results could be biased upwards in two ways. First, when units act as a treated group, any increases in spending that these grants induce will appear in the reduced form estimate of the effect of grant receipt on targeted spending, but the the associated first stage estimates will not capture increases in grant funding; second, the treatment effects of these grants will change over time (in particular, spending will return to pre-treatment levels as grant funding runs out), and when units act as controls, these changes in treatment effects will be subtracted out of our reduced form estimates.

1.6.4.2 Clean Watershed Needs Survey Data

We define whether a facility was in compliance with the CWA's capital mandate using the 1972 Clean Watershed Needs Survey, which we merge to our grant data with a unique facility code. The CWNS is an assessment of the capital investment that publicly-owned wastewater treatment facilities needed to make in order to come into compliance with the Clean Water Act and contains information on which community the facility serves, the total wastewater flowing through the facility, the treatment technology currently in place, whether the facility needs to meet standards higher than the EPA's secondary treatment requirement and whether they are currently in compliance with these requirements. This data comes from the EPA's CWNS team, and is the same data that Jerch (2018) uses to define compliance with the CWA's capital mandate.

We use a facility's answer to Question 21 on the CWNS questionnaire (reproduced in Figure 1.33) to define compliance. Question 21b asks if a facility needs to meet treatment technology requirements that are more stringent than the EPA's secondary treatment requirement. Question 21c then asks whether a facility is currently in compliance with both the EPAs secondary treatment mandate and any higher mandates. We define facilities that answered "yes" on question 21c as "compliant", and those that answer "no" as "non-compliant". Table 1.35 shows cross-tabs of these variables. Our sample includes facilities that were in compliance with the CWA's capital mandate when the CWA came into effect but were not in compliance with more stringent state standards. These facilities still faced incentives to spend CWA grants on sewerage capital, so we classify them as non-compliant.

Note that many facilities installed tertiary treatment after the CWA came into effect (USEPA, 2000). This increase was likely driven by municipalities bound by state standards or compelled by lawsuits to make upgrades beyond secondary treatment. Since we define these municipalities as non-compliant and we document an increase in sewerage capital spending in response to grant receipt in non-compliant municipalities, it is possible that municipalities used CWA grants to upgrade to tertiary treatment.

Our results are robust to dropping municipalities with facilities that were in compliance with the CWA's capital mandate but were not in compliance with more stringent state standards. Panels A and B of Table 1.37 re-estimate the results in Panels A and B of Table 1.4 after dropping non-compliant municipalities that had to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement. There is still evidence of full pass-through in municipalities that received grants smaller than $125 per capita and substantial crowd-out in municipalities that received grants larger than this amount.

We re-estimate the point where costs were no longer increasing in grant size using this sample in Table 1.38. This produces an estimate that is similar to the results in Table 1.3. We re-estimate pass-through for non-compliant facilities that were not required to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement that received grants above and below this point in panels C and D of Table 1.37. Again, the results are similar to those in Table 1.4.

Since our data on compliance is nearly 50 years old, we assessed its quality by comparing it to information on treatment technology from Table 2.4 of USEPA (2000). This table shows that, of 19,355 publicly owned wastewater treatment facilities surveyed, 9,887 were using secondary treatment or greater in 1972. This represents 51% of facilities. While we only observe whether a facility was in compliance with both EPA and state standards as opposed to just the secondary treatment technology standard, we can observe the percent of facilities that had at least secondary treatment among facilities that were only required to meet the EPA's standard. Column 1 of Table 1.35 shows that 56% of these facilities were using secondary treatment or greater in 1972. Given the different sample, this is relatively close to the 51% of facilities using secondary treatment or greater in 1972 in USEPA (2000).

Our definition of compliance does not directly enter into our pass-through equations. Instead, we use compliance to define sub-groups that we expect to find heterogeneous treatment effects for. If we mis-classify non-compliant municipalities as compliant, our estimates of pass-through for non-compliant municipalities will not be biased, but our estimates of pass-through for compliant municipalities would be biased upwards since they would include the effect of grant dollars that were bound by the CWA's capital mandate, and thus theoretically more likely to be spent on sewerage capital. Conversely, if we mis-classify compliant municipalities as non-compliant, our estimates of pass-through for compliant municipalities will not be biased, but our estimates of pass-through for non-compliant municipalities would be biased downwards since they would include the effect of more grant dollars that were not bound by the CWA's capital mandate.

Our definition of compliance does directly enter into our estimates of the effect of grant receipt on water revenue in Section 1.4. Mis-classifying municipalities as compliant or non-compliant would attenuate these results.

1.6.4.3 Estimating Costs

We define compliance in such a way that compliant municipalities never face any mandated costs, but compliance is not the only source of heterogeneity in mandated costs; there is also variation in mandated costs within non-compliant municipalities. The costs of upgrading depend largely on facility size (USEPA, 1973), and we use engineering estimates of the costs of upgrading from USEPA (1973) to approximate these costs.

USEPA (1973) uses data from the observed costs of upgrading plants to secondary treatment and fits this data to the number of gallons of wastewater flowing through a facility. This yields the following equation:

$$TotalCost = exp(.13732 + .77872 * ln(TotalFlow))$$

We use this equation to calculate mandated costs for each facility, then sum these costs across the facilities operated by a municipality to approximate the total costs that each municipality faces. While this measure of mandated costs is noisy, since it is based on an engineering formula, it is unlikely to be correlated with local preferences.[56]

These cost estimates never enter into any of our equations directly. Instead, we use them to determine if grants are larger or smaller than mandated costs. Using these cost estimates, we find that, while grants are similar to costs for municipalities that receive up to around \$125 per capita in total grant aid, grants to municipalities that receive more than this amount are substantially larger than the estimated cost of upgrading. We then separately estimate pass-through for municipalities that received grants above and below \$125 per capita in total grant aid. Noise in our cost estimates will cause us to mis-classify non-compliant municipalities between these two groups. This will attenuate our pass-trough estimates for non-complaint municipalities that received small grants, since it will cause these estimates to include non-compliant municipalities that received grants that included funding that was not bound by the capital mandate, and will bias our pass-trough estimates for non-complaint municipalities that received large grants upwards, since it will include

[56]If we used an assessment conducted by municipalities instead, we might be concerned that self-reported total costs are correlated with preferences for spending.

municipalities whose grants were completely bound by the capital mandate. Noise in our cost estimates will also cause our pass-through estimates for both groups to lose precision.

1.6.4.4 Merging Facility Data to Municipal Spending Data

The CWNS contains data on the authority that operates each facility, and we use this information to merge our facility-level data to municipality-level spending data from the Census Bureau's Historical Database on Individual Government Finances. The CWNS data does not contain a unique identifier of municipalities that is consistent with our local government finance data, but since we only have a balanced panel of spending data for 216 municipalities, it is relatively straightforward to use the state and municipality name in both detests to match our municipality level data to the facilities that they operated.

1.6.4.5 Measuring Spending in Real Per Capita Dollars

We follow the previous pass-through literature by measuring grant amount and all spending in per capita terms. (Strumpf, 1998; Card and Payne, 2002; Vegh and Vuletin, 2015). All of our variables of interest are largely influenced by population, but dividing these variables by population ensures that our estimates are not biased by population growth. Figure 1.34 re-estimates the results in Figure 1.4 with population as the dependent variable, and shows that population did not change after grant receipt.

We deflate grant amount and all spending variables to 1973 dollars with the Consumer Price Index. As shown in Figures 1.35, 1.36 and 1.37, and Tables 1.39, 1.40 and 1.41, our main results are robust to measuring spending and grant amount in nominal terms.

1.6.5 Additional Figures and Tables

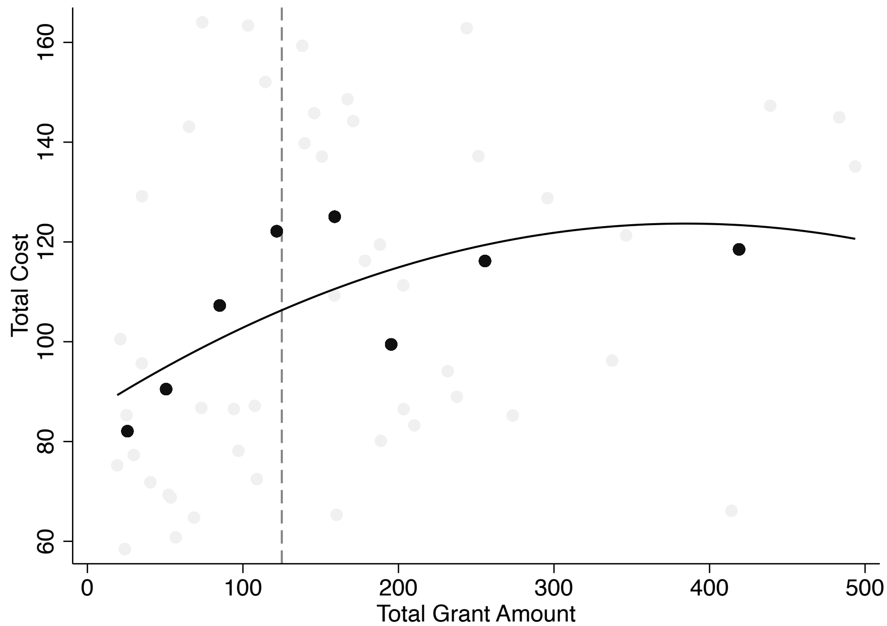


Figure 1.8: Cost Data with Quadratic Fit

Notes: This figure re-creates Figure 1.1 a quadratic fit.

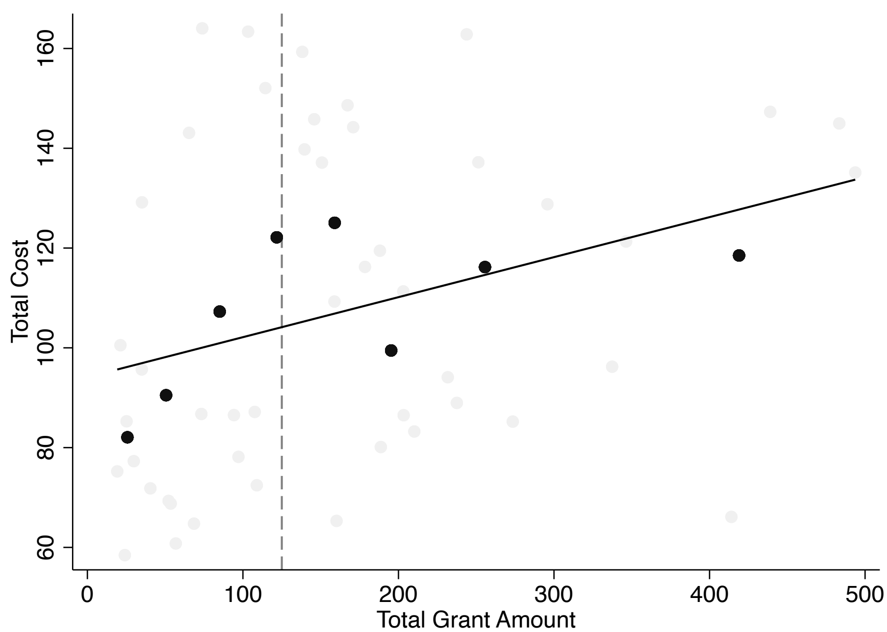


Figure 1.9: Cost Data with Linear Fit

Notes: This figure re-creates Figure 1.1 with a a linear fit.

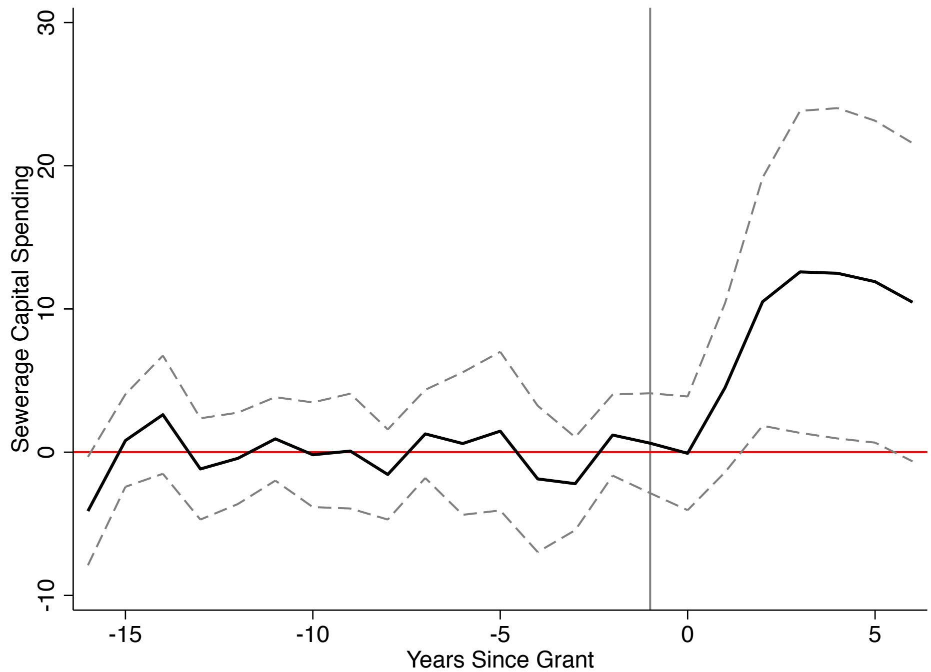


Figure 1.10: Balanced Sewerage Capital Event Study

Notes: This figure re-estimates the results in Figure 1.4 on a sample of municipalities that we observe for at least seven post-treatment periods.

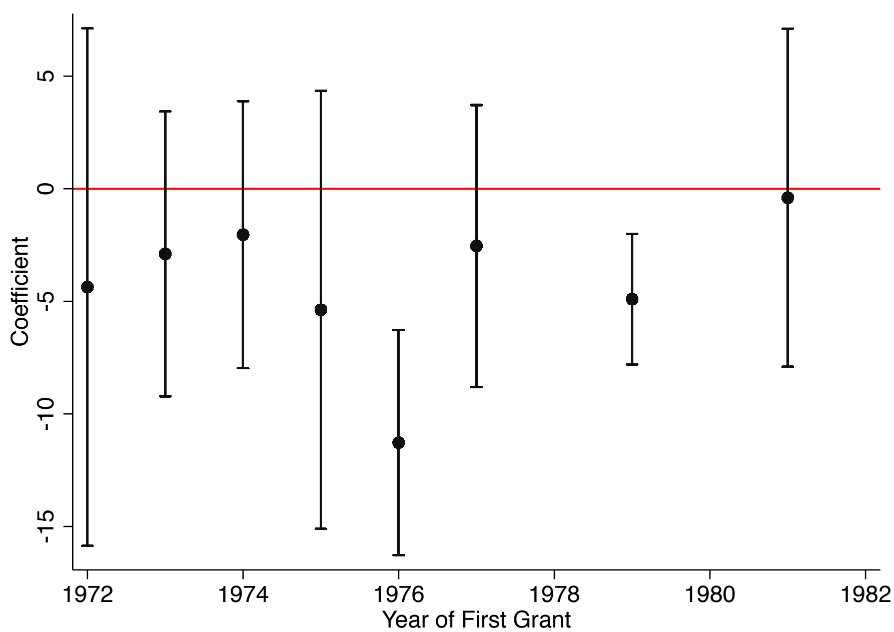


Figure 1.11: Decomposition of Main Water Revenue Estimate

Notes: This figure presents estimates of the θ_g^{2X2} from estimating $R_{it} = \alpha_0 + \theta_g^{2X2} 1\{t \geq g\} * compliant_i + \delta 1\{t \geq g\} + \omega compliant_i + \varepsilon_{it}$ on observations from each timing group g along with the 95% confidence interval of each θ_g^{2X2}. We plot each θ_g^{2X2} against the year timing group g became treated.

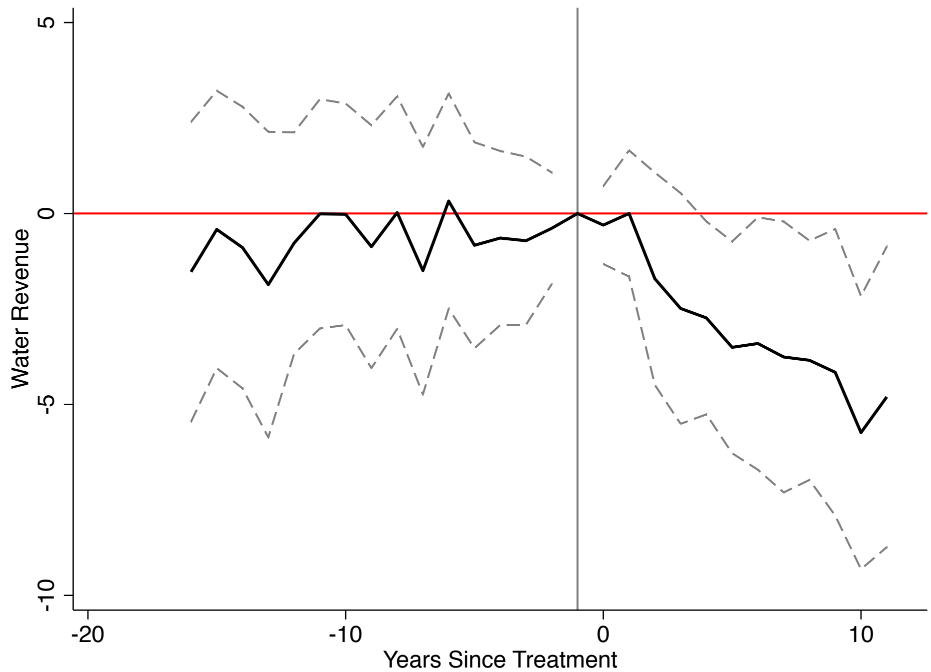


Figure 1.12: Water Revenue Event Study with Region-by-Year Fixed Effects

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_{gt} + \alpha_{rt} + \varepsilon_{it}$ with per water revenue as the dependent variable. $compliant_i$ equals one for observations of compliant municipalities. We include municipality, timing-group-by-year, and region-by-year fixed effects, α_i, α_{gt}, and α_{rt}, where g indexes the year in which a municipality received its first CWA grant and r indexes region.

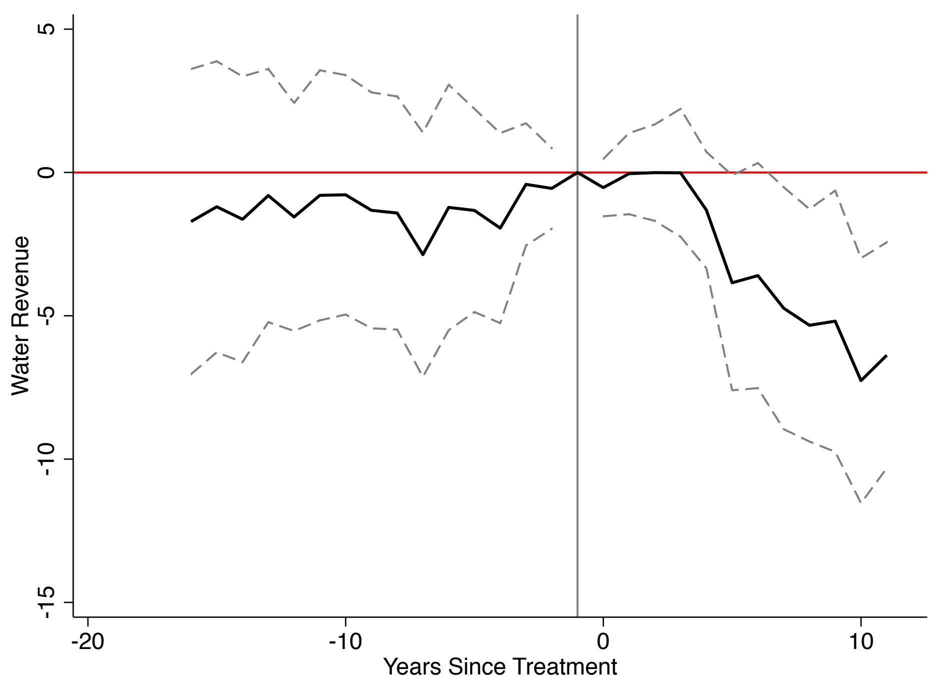


Figure 1.13: TWFE Water Revenue Event Study

Notes: This figure plots the π_y and γ_y from estimating $R_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * compliant_i + \sum_{y=0}^{11} \gamma_y 1\{t - t_i^* = y\} * compliant_i + \alpha_i + \alpha_t + \varepsilon_{it}$. α_i and α_t are municipality and year fixed effects.

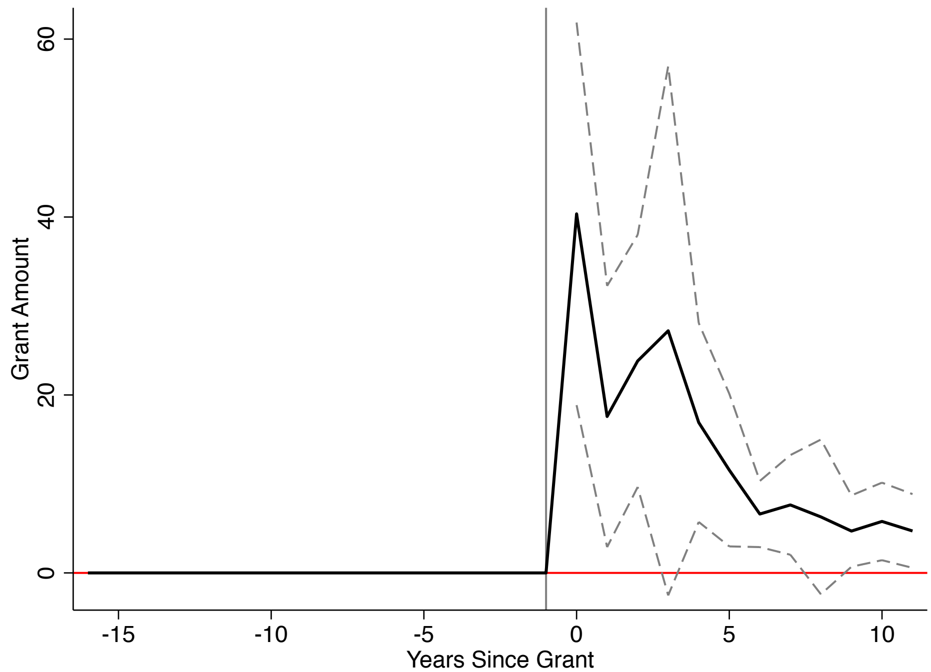


Figure 1.14: Grant Amount Event Study Using Never-Treated Controls

Notes: This figure re-creates Figure 1.3 using never-treated units as a control group.

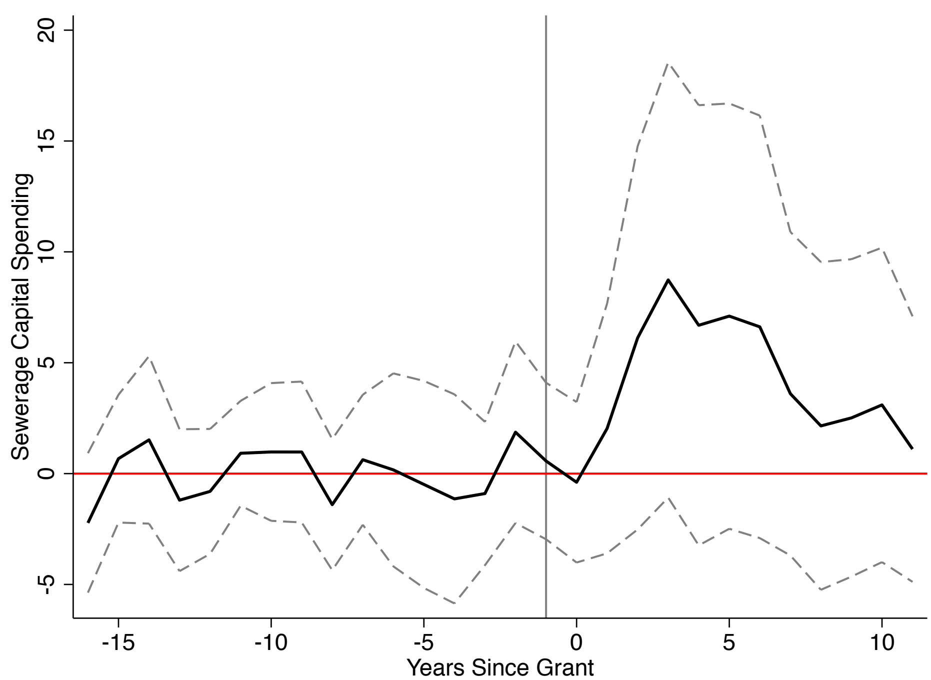


Figure 1.15: Sewerage Capital Spending Event Study Using Never-Treated Controls

Notes: This figure re-creates Figure 1.4 using never-treated units as a control group.

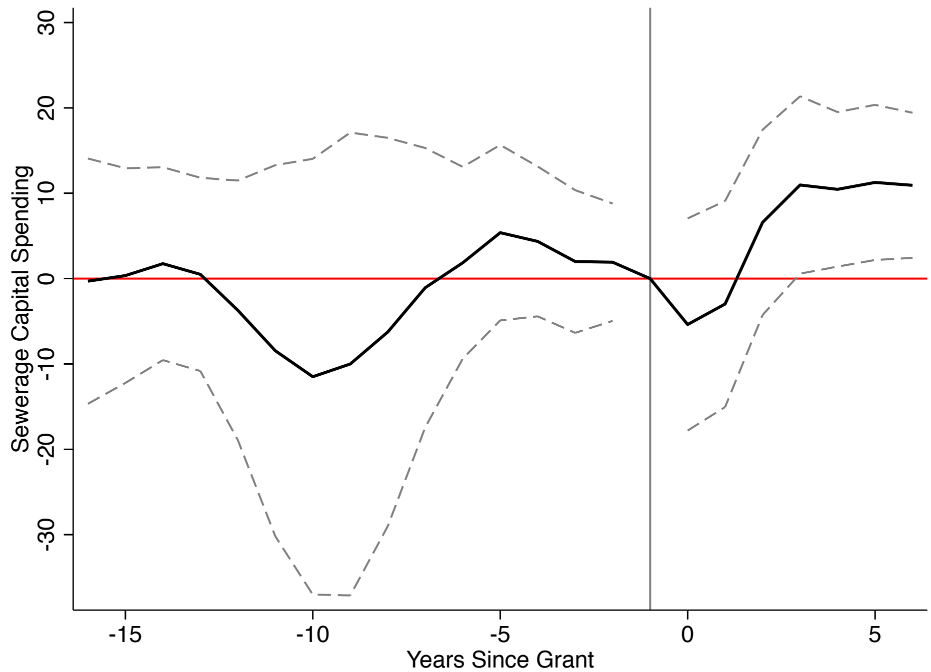


Figure 1.16: Stacked Sewerage Capital Event Study

Notes: This figure plots the π_y and γ_y from estimating $C_{it} = \alpha_0 + \sum_{y=-16}^{-2} \pi_y 1\{t - t_i^* = y\} * treat_{is} + \sum_{y=0}^{6} \gamma_y 1\{t - t_i^* = y\} * treat_{is} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$. i indexes municipalities, t indexes years and s indexes stacks. $treat_{is}$ equals one for observations of the treated municipality in each stack after that municipality received its first CWA grant. The dependent variable is per capita sewerage capital spending and α_{is} and α_{ts} are stack by municipality and stack by year fixed effects.

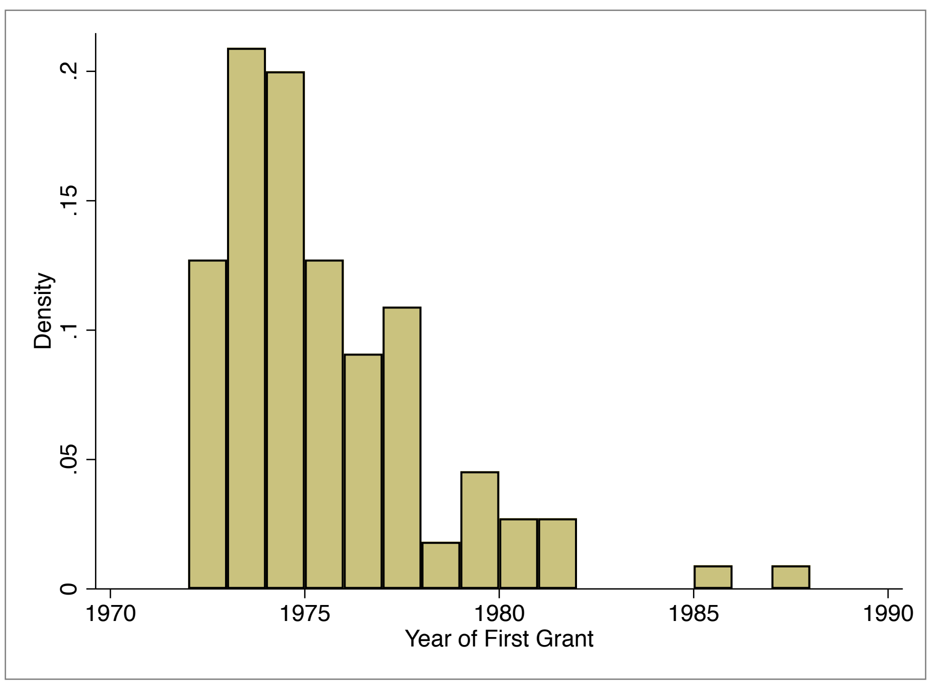


Figure 1.17: Timing Groups

Notes: This figure shows how many municipalities received their first CWA grant in each year from 1972 to 1988.

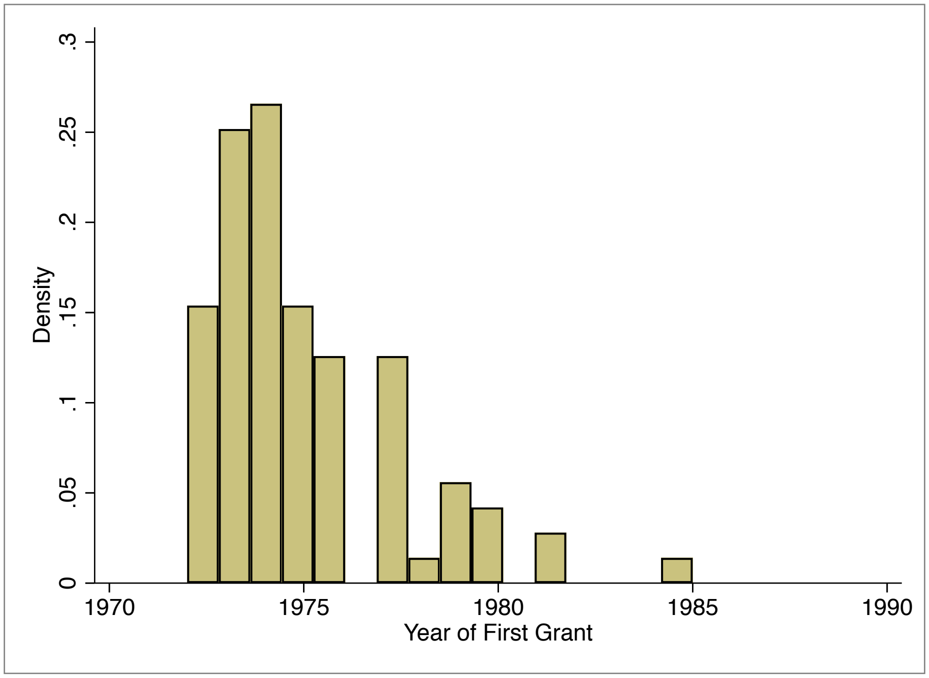


Figure 1.18: Non-Compliant

Notes: This figure re-creates Figure 1.17 for non-compliant municipalities.

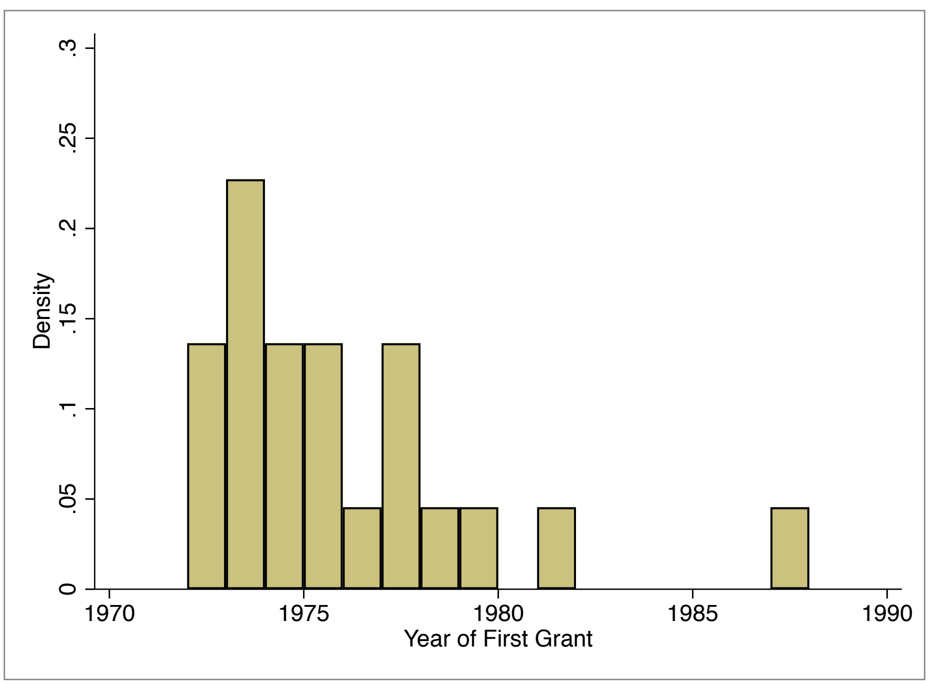


Figure 1.19: Timing Groups for Compliant Municipalities

Notes: This figure re-creates Figure 1.17 for compliant municipalities.

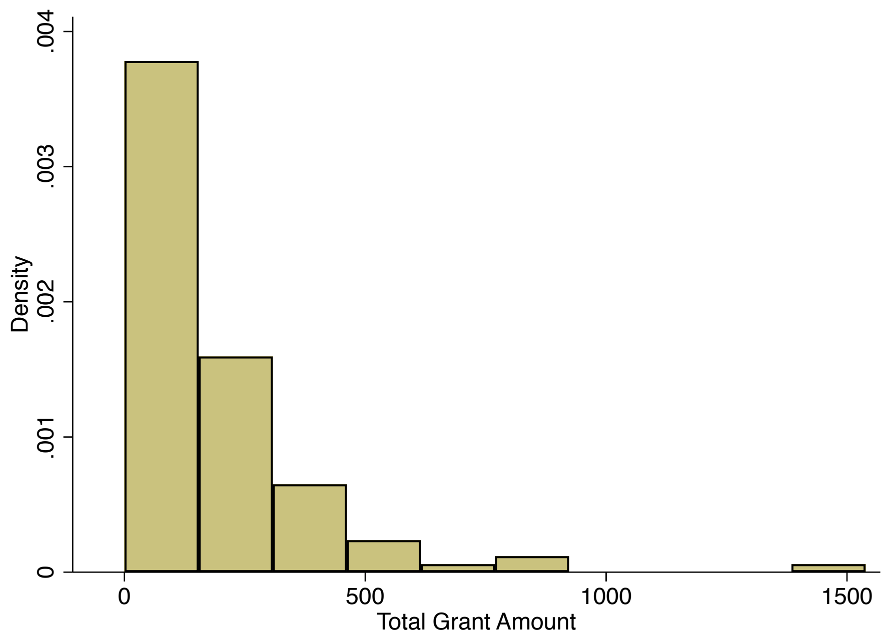


Figure 1.20: Distribution of Grant Size

Notes: This figure shows the distribution of total grant dollars.

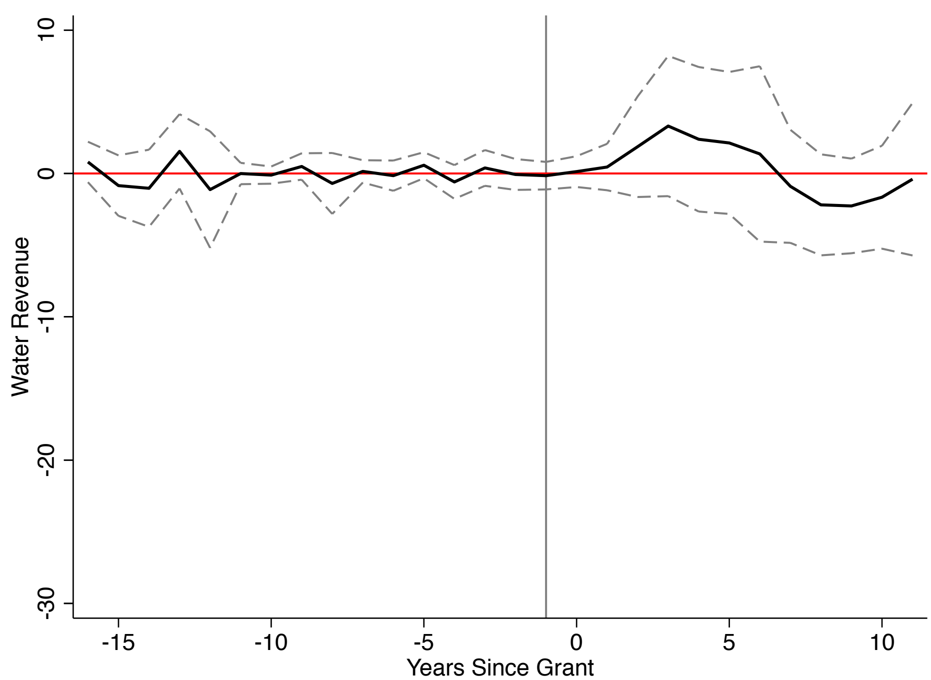


Figure 1.21: Water Revenue Did Not Change After Grant Receipt in Non-Compliant Municipalities

Notes: This figure presents $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G = g|t-g=e)$ for non-compliant municipalities. Water revenue is the dependent variable.

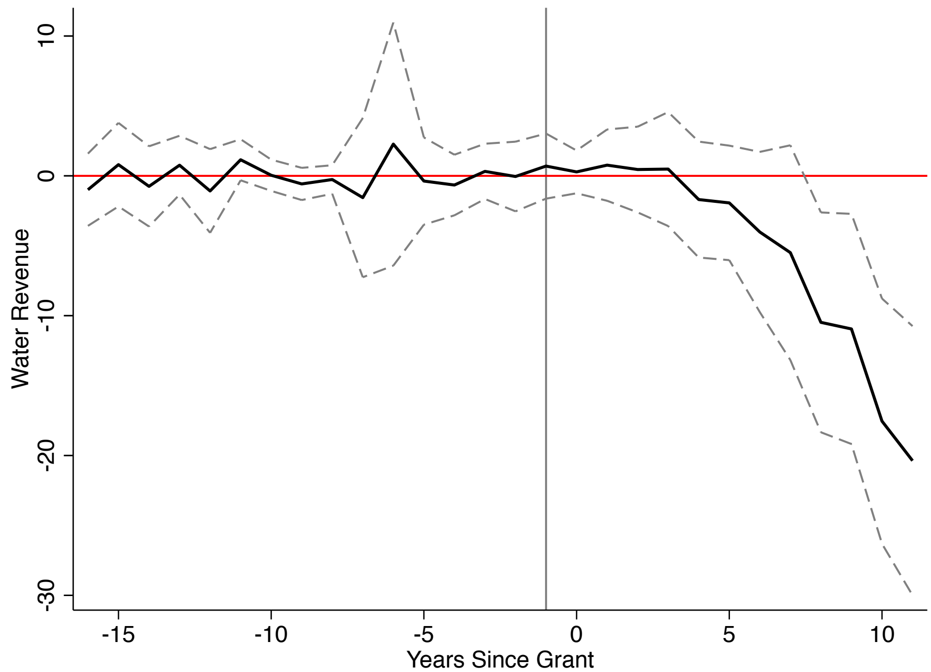


Figure 1.22: Water Revenue Decreased After Grant Receipt in Compliant Municipalities

Notes: This figure re-creates Figure 1.21 for compliant municipalities

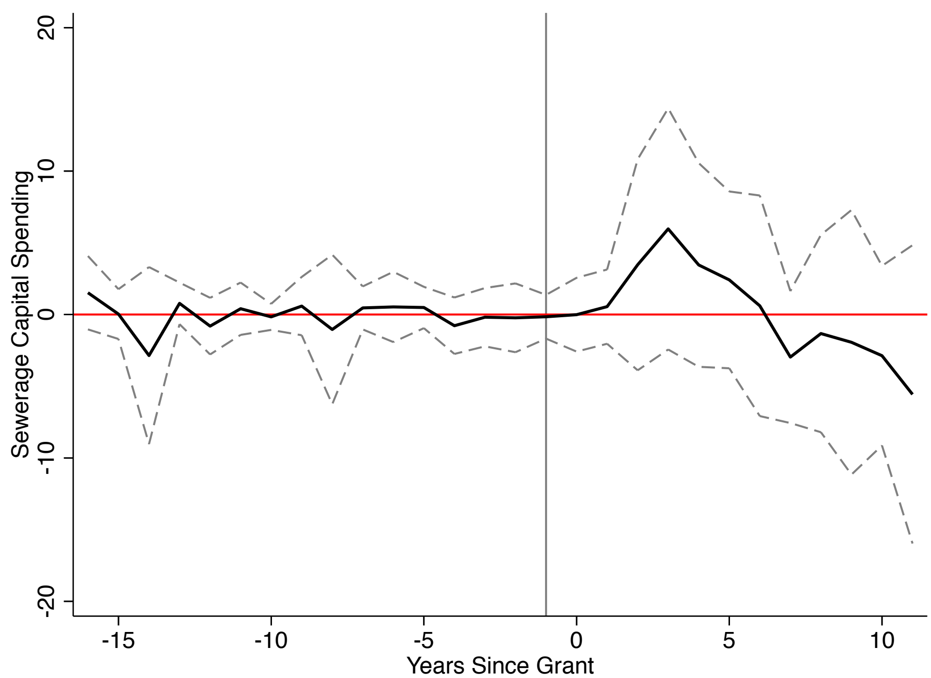


Figure 1.23: Non-Complaint Municipalities Raised Water Bills When Grants Were Too Small

Notes: This figure re-estimates the results in Figure 1.21 on sub-samples of municipalities where grants are smaller than estimated costs.

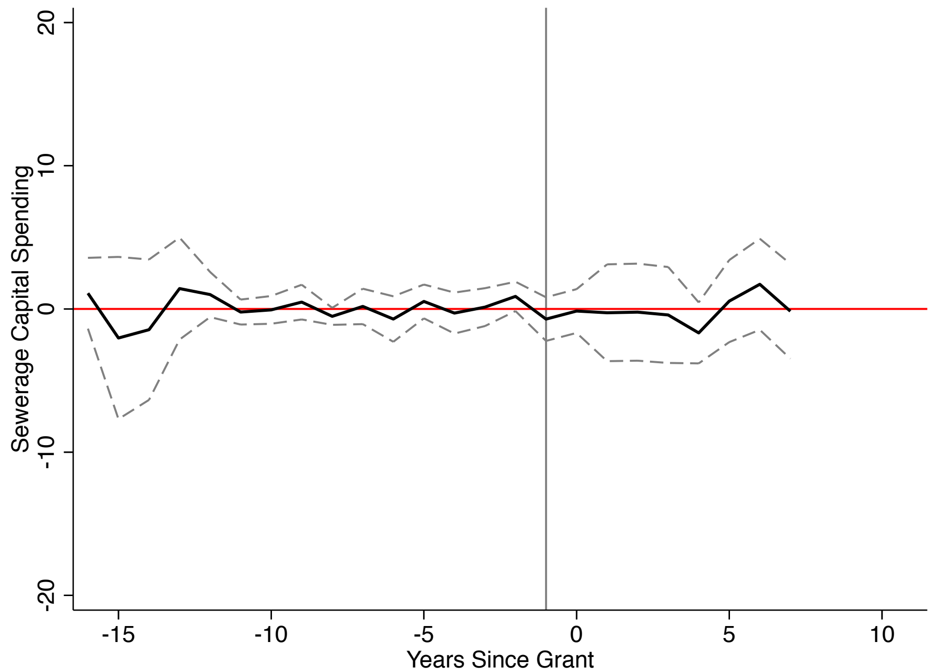


Figure 1.24: Non-Complaint Municipalities Did Not Change Water Bills in Response to Large Grants

Notes: This figure re-estimates the results in Figure 1.21 on sub-samples of municipalities where grants are larger than estimated costs.

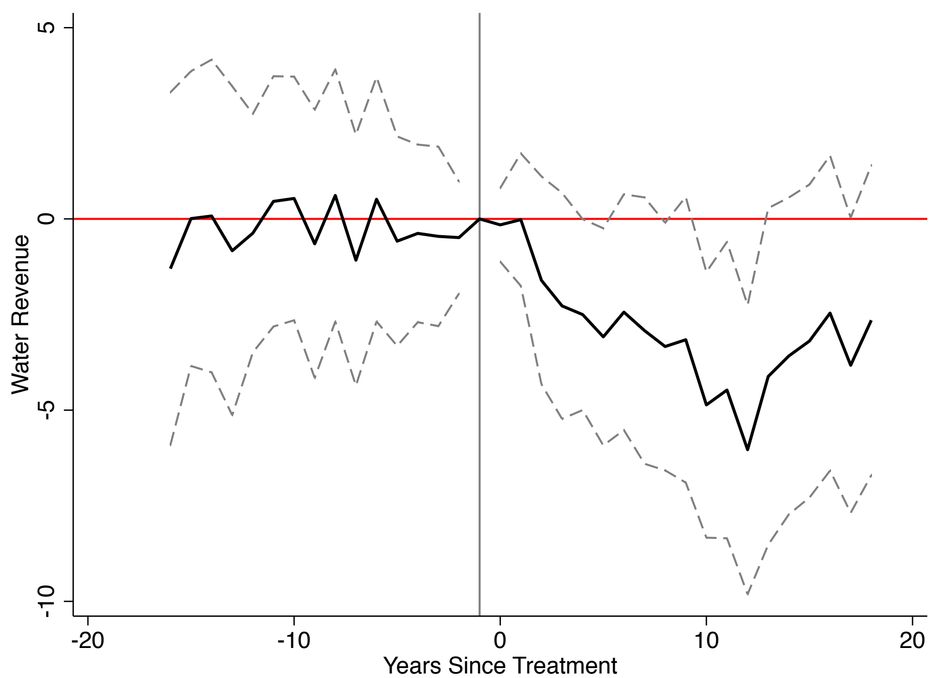


Figure 1.25: The Effect of Grants on Water Revenue Flattens Out

Notes: This figure re-estimates the results in Figure 1.7 on municipalities treated before 1982.

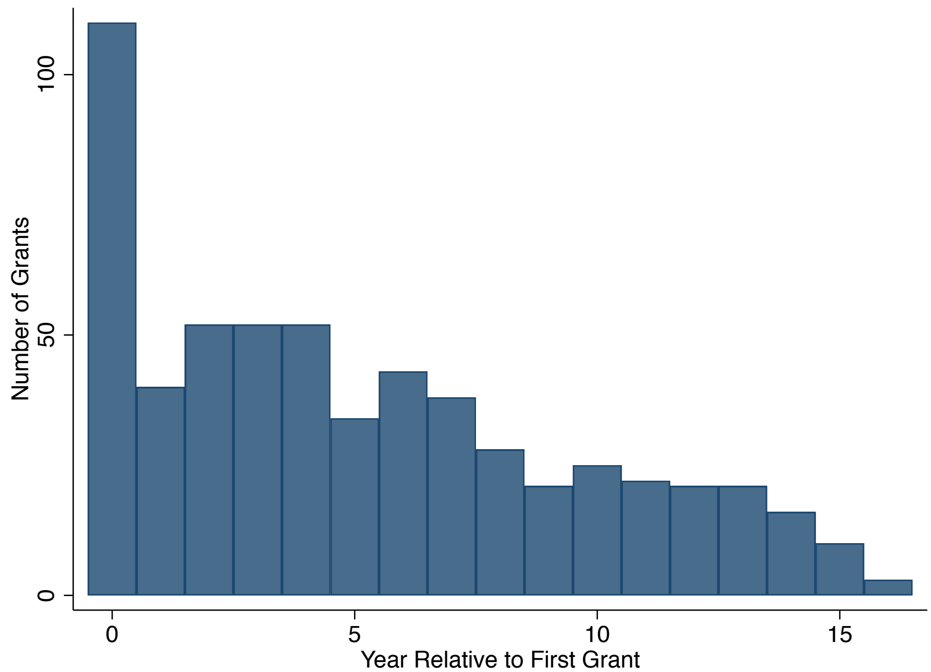


Figure 1.26: Distribution of Grants Over Time

Notes: This figure shows the number of grants that municipalities received each year relative to the year of the municipality's first grant.

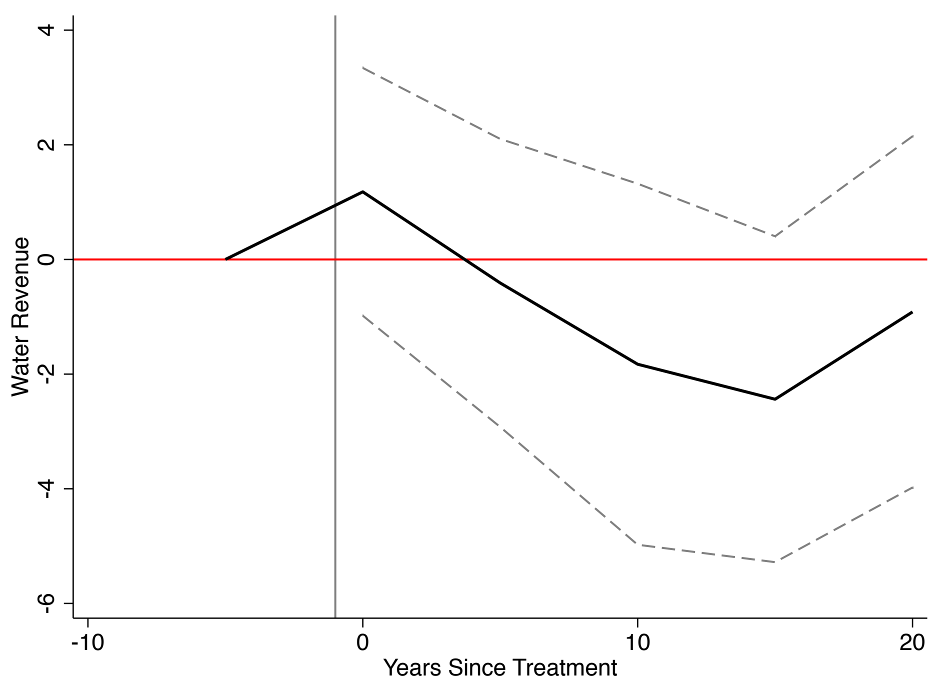


Figure 1.27: Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 1.7 on an alternative dataset.

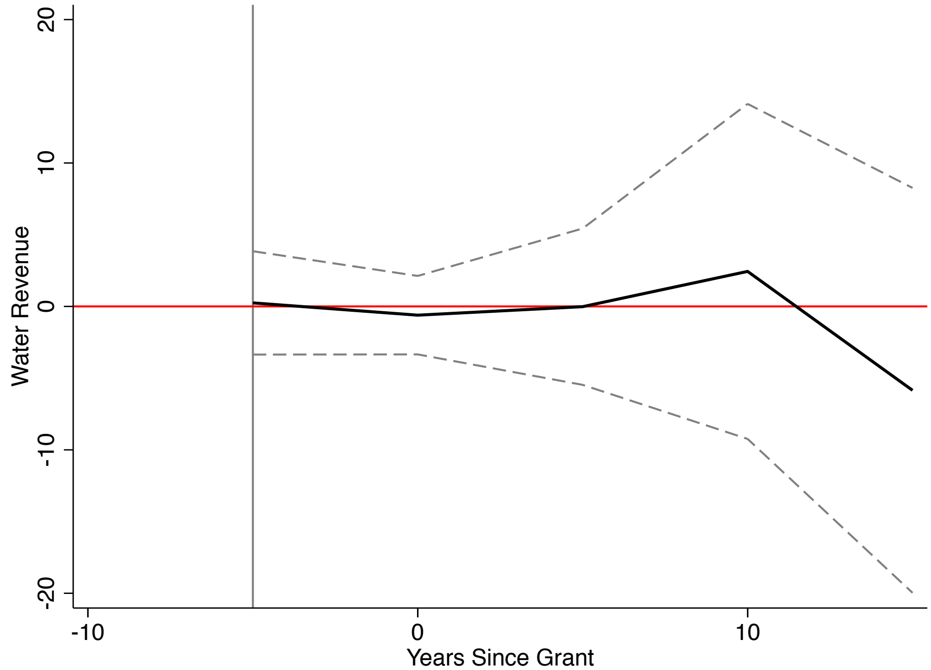


Figure 1.28: Non-Compliant Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 1.21 on an alternative dataset.

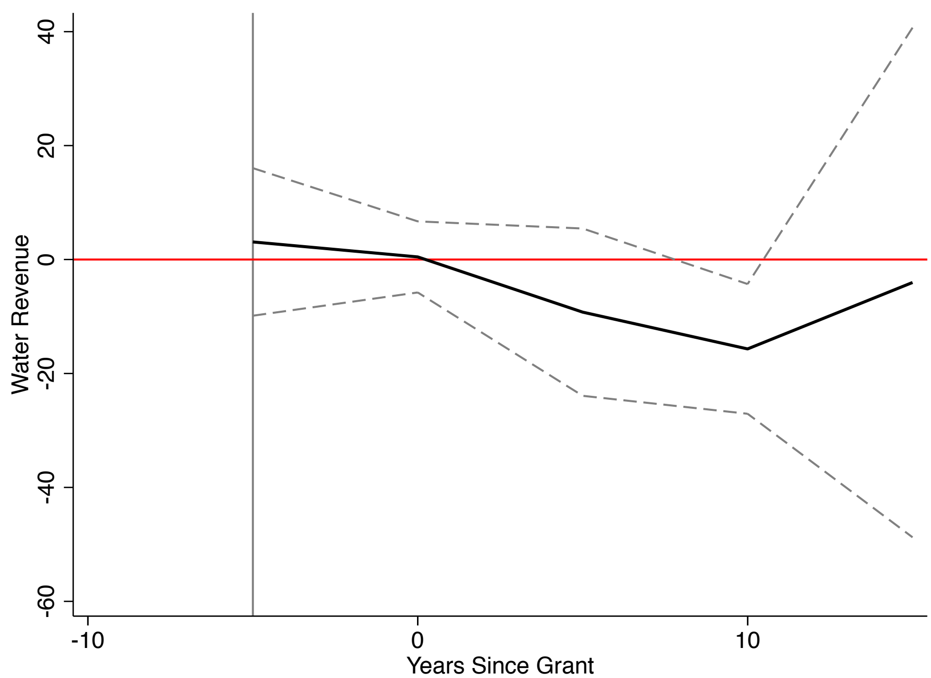


Figure 1.29: Compliant Water Revenue Event Study with Alternative Data

Notes: This figure re-estimates the results from Figure 1.22 on an alternative dataset.

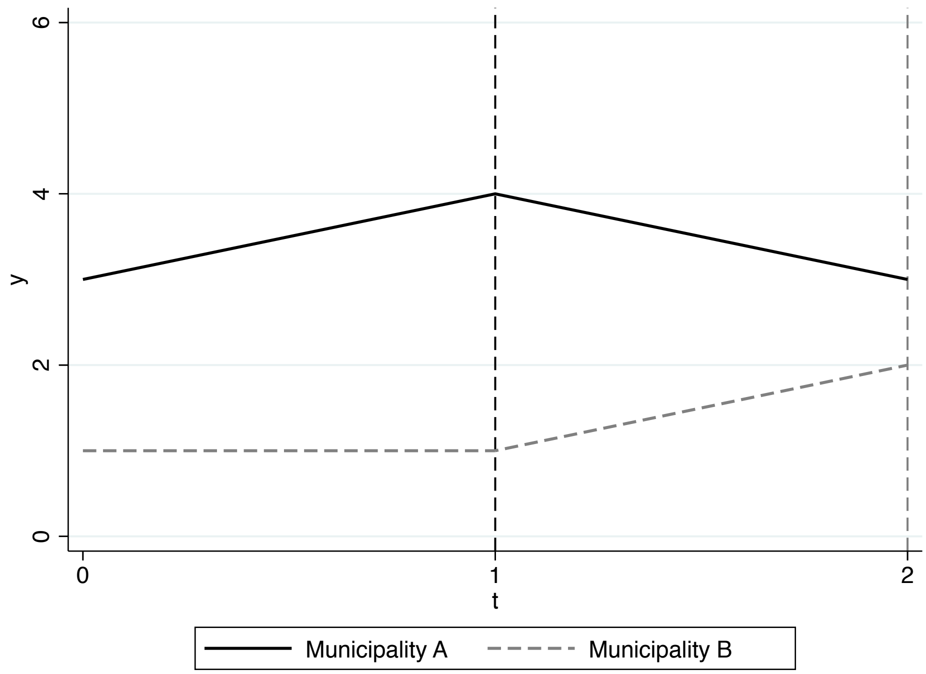


Figure 1.30: Example of Problems with Binary TWFE

Notes: This figure shows a hypothetical situations where two municipalities are treated at different times. Municipality A is treated in period 1 and Municipality B is treated in period 2.

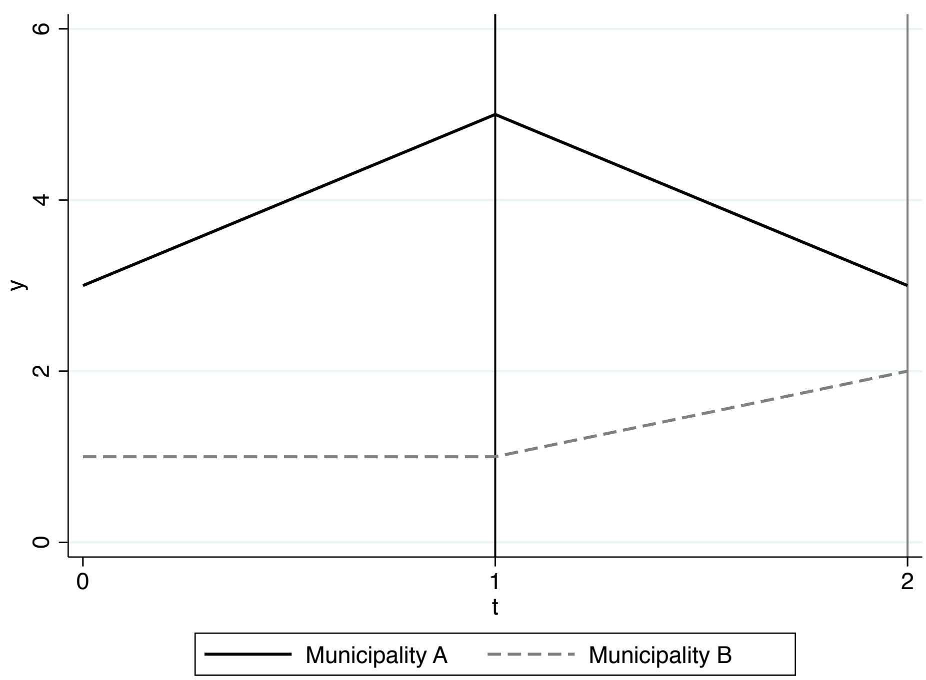


Figure 1.31: Example of Problems with Continuous TWFE

Notes: This figure shows another hypothetical situations where two municipalities are treated at different times.

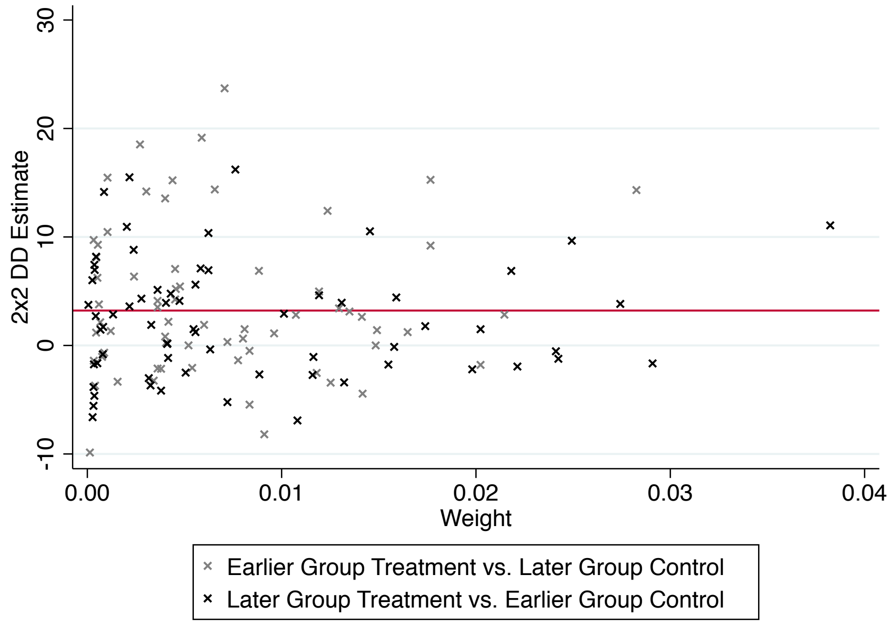


Figure 1.32: Bacon Decomposition of Sewerage Capital Spending TWFE Estimate

Notes: This figure presents a scatterplot of all possible two-group/two period DD estimators and their associated weights. These estimates and weights come from a decomposition of β_{rf}^{TWFE} from $C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ made using Goodman-Bacon et al. (2019). C_{it} is per capita sewerage capital expenditure, D_{it} is a dummy variable that equals one after grant receipt and α_i and α_t are municipality and year fixed effects.

Authority and Facility No. ▶ __ __ - __ __ __ __ - __ __ __

Section III – CATEGORY I – ASSESSMENT OF NEEDS BY TYPE AND COST

ASSESSMENT OF NEEDS TO ACHIEVE REQUIRED LEVEL OF SECONDARY TREATMENT	*Mark appropriate box for each item*	**EPA USE ONLY**
21a. Can this plant meet water quality standards applicable to the stream segment to which it discharges by a level of treatment LESS than defined as secondary treatment by EPA?	1 ☐ Yes 2 ☐ No 3 ☐ Not known 4 ☐ Not applicable; no discharge to waters	
b. What level of secondary treatment must the discharge from this plant meet by July 1, 1977?	1 ☐ Secondary treatment level as defined by EPA, **OR** 2 ☐ Higher level of secondary treatment required by State – *Specify* Higher level secondary treatment Nature of State action	
c. Does the discharge from this plant **NOW** meet the level of secondary treatment identified in 21b?	1 ☐ Yes 2 ☐ No	
d. Will the discharge from this plant meet on July 1, 1977, the level of secondary treatment identified in 21b? *(Give consideration to changes in flow and concentration of influents and to the changes in treatment capability now under construction or provided for in approved grants.)*	1 ☐ Yes – *SKIP to item 25* 2 ☐ No	
e. Which approaches will be used to enable this existing or proposed plant to meet the secondary treatment level identified in 21b? (1) Addition of land disposal as a means of treatment	1 ☐ Yes 2 ☐ No	
(2) New plant – no replacement of existing plant	1 ☐ Yes 2 ☐ No	
(3) Replacement plant	1 ☐ Yes 2 ☐ No	
(4) Modification – no change in capacity or treatment level	1 ☐ Yes 2 ☐ No	
(5) Modification – increase in capacity	1 ☐ Yes 2 ☐ No	
(6) Modification – increase in treatment level	1 ☐ Yes 2 ☐ No	
(7) Improved operation and maintenance, increase staffing	1 ☐ Yes 2 ☐ No	
(8) Reduce infiltration	1 ☐ Yes 2 ☐ No	

Page 15

Figure 1.33: CWNS Compliance Question

Notes: This figure reproduces the question from the 1972 CWNS that we use to define compliance.

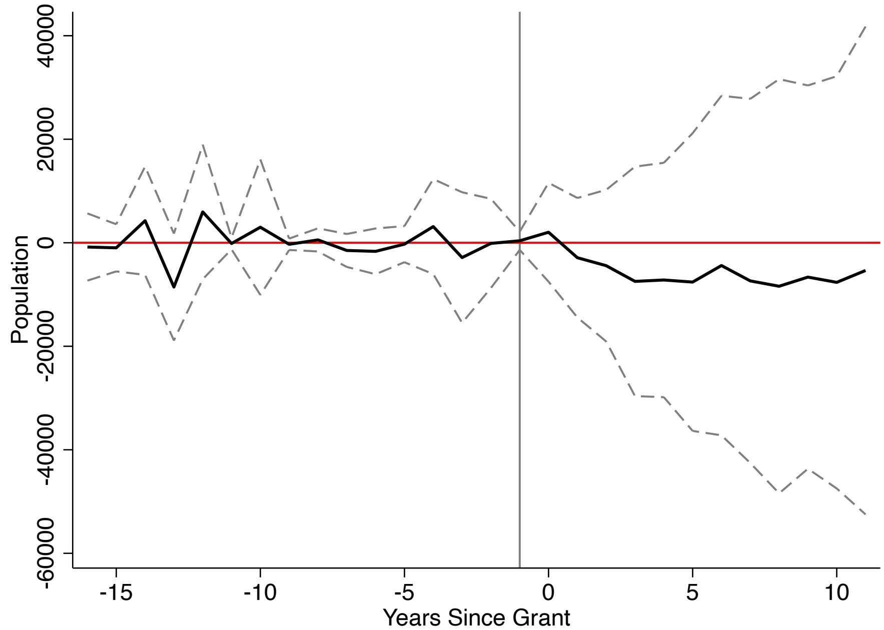


Figure 1.34: Population Did Not Change After Grant Receipt

Notes: This figure shows the reduced form relationship between grant receipt and population by presenting the $\widehat{\theta}_D(e)$ for each year e relative to treatment, where $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985} \sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e)$. g indexes timing groups and t indexes years. $\widehat{ATT(g,t)}$ is the difference between changes in population in units in group g in periods t and $g-1$ and changes in population in units not yet treated at time t in periods t and $g-1$.

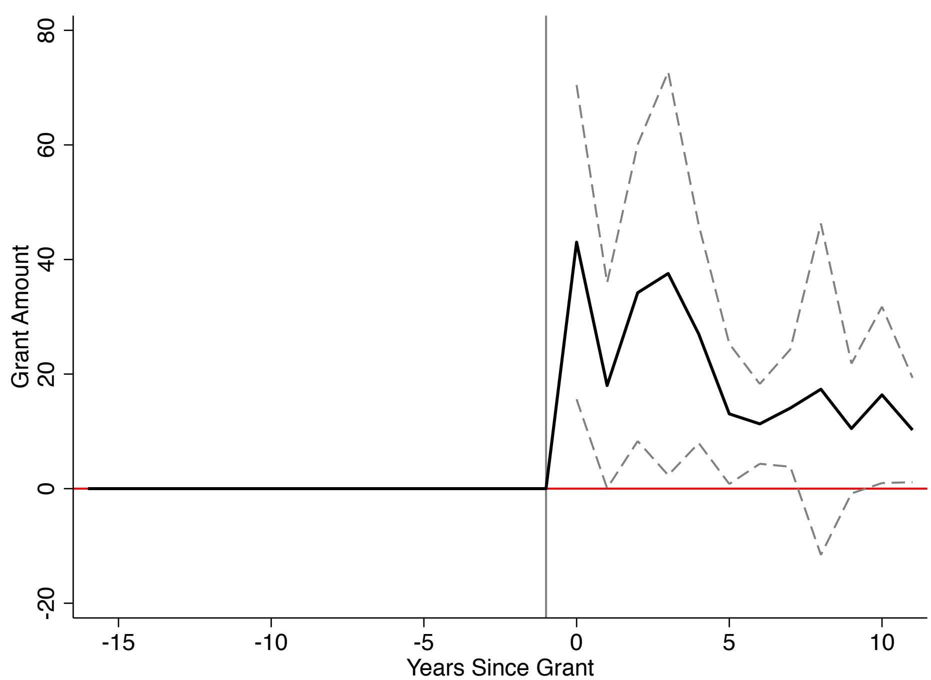


Figure 1.35: Grant Amount Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 1.3 measuring grant amount in nominal dollars.

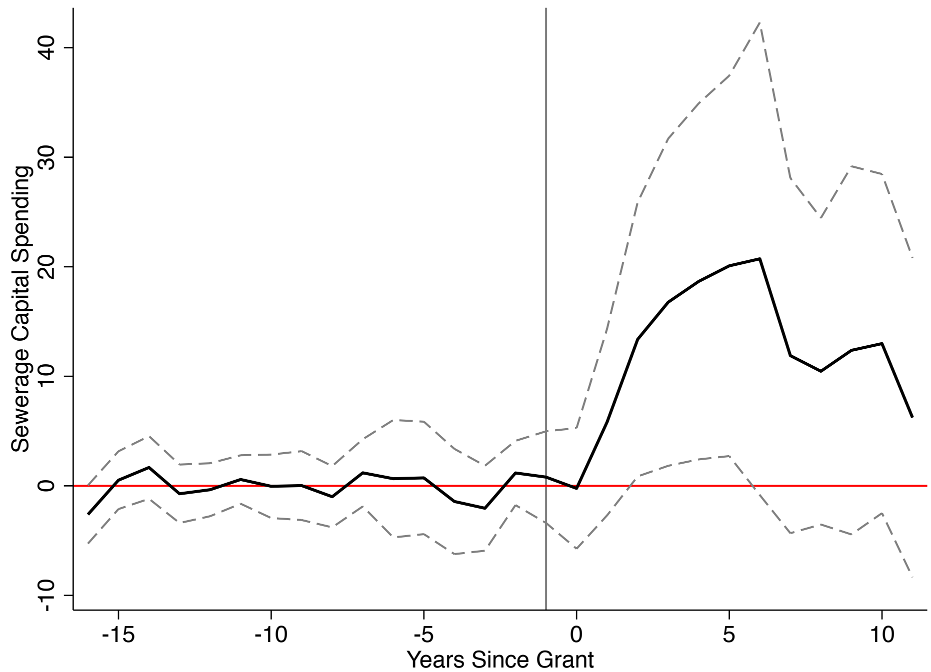


Figure 1.36: Sewerage Capital Spending Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 1.4 measuring sewerage capital spending in nominal dollars

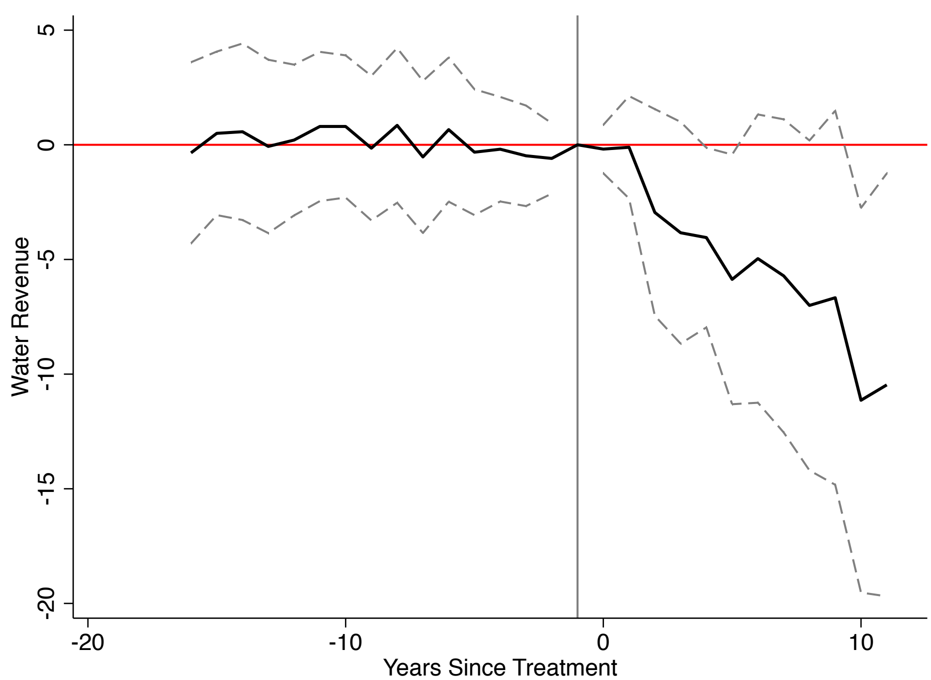


Figure 1.37: Water Revenue Event Study (Nominal Dollars)

Notes: This figure re-creates Figure 1.7 measuring water revenue in nominal dollars.

Table 1.10: 1970 Summary Statistics by Facility Compliance and Grant Size

	(1) Compliant	(2) NC & < $125	(3) NC & >= $125
Population	115,501	398,565	395,286
	(73,126)	(722,518)	(1077607)
Total Revenue	253.13	290.61	296.40
	(138.17)	(136.48)	(166.63)
Total IGR	52.84	50.67	72.72
	(62.93)	(40.28)	(82.90)
Federal IGR	11.43	9.14	11.89
	(15.64)	(8.89)	(12.89)
Nonwater Utility Rev	21.18	32.53	9.42
	(44.79)	(60.27)	(30.16)
Property Tax Revenue	82.21	99.47	97.59
	(72.75)	(94.32)	(69.14)
Water Utility Revenue	23.10	22.93	22.90
	(9.10)	(8.84)	(8.50)
Total Expenditure	277.26	302.29	311.73
	(154.56)	(145.95)	(179.20)
Total IG Exp	1.93	4.79	4.38
	(4.54)	(9.98)	(10.07)
Capital Outlays	79.35	68.87	60.84
	(57.33)	(41.85)	(34.90)
Highway Exp	21.16	21.25	22.37
	(8.38)	(9.54)	(9.38)
Sewerage Capital Outlay	8.94	6.66	5.87
	(9.61)	(6.50)	(7.01)
Sewerage Non-capital Exp	5.50	4.91	5.12
	(3.13)	(3.37)	(2.97)
Nonwater Utiltiy Exp	22.46	31.23	11.28
	(49.25)	(56.94)	(38.13)
Water IG Exp	1.98	2.91	2.07
	(1.91)	(3.22)	(2.09)
Water Non-capital Exp	13.34	12.03	13.35
	(4.74)	(7.16)	(4.99)
Water Capital Outlay	13.45	11.84	6.27
	(19.01)	(16.15)	(6.83)
Observations	22	34	56

Notes: This table presents summary statistics for municipalities with compliant facilities, non-compliant facilities that received small grants, and non-complaint facilities that received large grants in 1970. All covariates aside from the facility compliance dummy and population represent per capita values.

Table 1.11: Observable Characteristics Do Not Predict Grant Size or Compliance

	(1) Grant >= 125	(2) Compliant
ln Population	0.0315	-0.0662*
	(0.0595)	(0.0370)
Total Revenue	-0.00121	-0.00276
	(0.00299)	(0.00193)
Total IGR	0.00311*	-0.000142
	(0.00180)	(0.00140)
Federal IGR	0.00159	0.000359
	(0.00268)	(0.00194)
Nonwater Utility Rev	-0.00371	-0.00381
	(0.00288)	(0.00236)
Property Tax Revenue	-0.00127	0.000757
	(0.00118)	(0.000899)
Total Expenditure	0.00160	0.00162
	(0.00318)	(0.00211)
Total IG Exp	0.00249	-0.00349
	(0.00449)	(0.00358)
Capital Outlays	-0.00398	0.000949
	(0.00274)	(0.00207)
Highway Exp	-0.00155	0.00177
	(0.00503)	(0.00428)
Sewerage Capital Outlay	0.00267	-0.00294
	(0.00509)	(0.00351)
Sewerage Non-capital Exp	-0.00203	0.00719
	(0.0147)	(0.0133)
Nonwater Utiltiy Exp	0.00157	0.00500*
	(0.00263)	(0.00276)
Water IG Exp	-0.0128	-0.0188
	(0.0281)	(0.0230)
Water Non-capital Exp	0.0269*	-0.00377
	(0.0160)	(0.00964)
Water Capital Outlay	-0.00158	0.00812
	(0.00534)	(0.00570)
Water Utility Revenue	-0.00519	-0.00155
	(0.0102)	(0.00782)
Observations	112	112

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents results from regressing dummies indicating municipalities that received grants totaling more than \$125 (in column 1) and compliant municipalities (in column 2) on potential predictive covariates from 1971. All covariates aside from the natural log of population represent per capita values.

Table 1.12: Event Study Coefficients

	Grant Amount	Sewerage Capital
e = -16	0.0000	-4.0571
		(1.3551)
e = -15	0.0000	0.8037
		(1.2730)

e = -14	0.0000	2.5714
		(1.4952)
e = -13	0.0000	-1.1177
		(1.2199)
e = -12	0.0000	-0.4861
		(1.2813)
e = -11	0.0000	0.9238
		(1.0574)
e = -10	0.0000	-0.1678
		(1.3288)
e = -9	0.0000	0.0719
		(1.4024)
e = -8	0.0000	-1.4882
		(1.0897)
e = -7	0.0000	1.2816
		(1.0761)
e = -6	0.0000	0.5743
		(1.7228)
e = -5	0.0000	1.4460
		(1.8526)
e = -4	0.0000	-1.7757
		(1.6682)
e = -3	0.0000	-2.1195
		(1.2274)
e = -2	0.0000	1.1961
		(0.9192)
e = -1	0.0000	0.6774
		(1.3010)
e = 0	39.8180	-0.2082
	(9.5342)	(1.3512)
e = 1	15.3706	4.3462
	(6.1208)	(2.1316)
e = 2	24.4435	10.3150
	(6.2881)	(3.5160)
e = 3	27.4015	12.5877
	(9.7786)	(3.6743)
e = 4	17.6683	12.4909
	(4.9166)	(3.9238)
e = 5	7.8658	11.9075
	(2.5698)	(3.8063)
e = 6	6.0865	10.4708
	(1.3436)	(4.0569)
e = 7	6.8891	4.5029
	(4.6979)	(3.2382)
e = 8	7.8224	3.3674
	(4.6979)	(2.8253)
e = 9	4.5610	3.1520
	(1.7787)	(2.7576)
e = 10	6.9069	3.2678
	(2.3469)	(3.0731)
e = 11	4.3978	0.3406
	(1.4327)	(2.4918)

Notes: This table presents the $\widehat{\theta}_D(e)$ from Figures 1.3 and 1.4.

Table 1.13: Pass-Through Results are Robust to Different Cutoffs

Panel A: Non-Compliant and < $138	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	7.396***	6.682	
	(.752)	(7.729)	
Pass-through			.904
			(1.047)
p-value: Pass-through = 0			0.388
p-value: Pass-through = 1			0.924
Observations	2254	2254	2254
Panel B: Non-Compliant and >= $138			
Treat * Post	47.80***	13.043	
	(6.788)	(7.337)	
Pass-through			0.273
			(0.152)
p-value: Pass-through = 0			0.073
p-value: Pass-through = 1			0.000
Observations	1426	1426	1426

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 1.4 using a different cutoff. Panel A summarizes pass-through for stacks defined by non-compliant municipalities who received grants totalling less than $138, and Panel B summarizes pass-through for all other stacks defined by non-compliant municipalities.

Table 1.14: Compliant Pass-Through is Low for Large and Small Grants

Panel A: Compliant and < $125	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	6.788***	-2.686	
	(1.825)	(1.557)	
Pass-through			-0.395
			(0.275)
p-value: Pass-through = 0			0.151
p-value: Pass-through = 1			0.000
Observations	598	598	598
Panel B: Compliant and >= $125			
Treat * Post	34.34***	6.567	
	(3.782)	(9.208)	
Pass-through			0.191
			(0.277)
p-value: Pass-through = 0			0.491
p-value: Pass-through = 1			0.004
Observations	322	322	322

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 1.4 on compliant municipalities.

Table 1.15: Pass-Through by Population Tercile

Panel A: Tercile 1	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	23.96***	8.551*	
	(8.208)	(4.596)	
Pass-through			0.357
			(0.240)
p-value: Pass-through = 0			0.137
p-value: Pass-through = 1			0.007
Observations	1584	1584	1584
Panel B: Tercile 2			
Treat * Post	18.08***	10.93***	
	(3.143)	(3.802)	
Pass-through			0.605**
			(0.279)
p-value: Pass-through = 0			0.030
p-value: Pass-through = 1			0.156
Observations	1628	1628	1628
Panel C: Tercile 3			
Treat * Post	20.58***	5.383**	
	(4.892)	(2.203)	
Pass-through			0.262
			(0.169)
p-value: Pass-through = 0			0.119
p-value: Pass-through = 1			0.000
Observations	1628	1628	1628

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results from Panel B of Table 1.2 by terciles of population. Tercile 1 has the lowest population in our sample, and Tercile 3 has the highest.

Table 1.16: Re-Weight Stacked Results to Reflect All CWA Grants

	(1) Non-Compliant	(2) Compliant	(3) Full Sample
Pass-Through	0.397	0.0337	0.353
	(0.310)	(0.363)	(0.278)
p value: Pass-through = 0	0.052	0.075	0.204
p value: Pass-through = 1	0.200	0.331	0.020
Observations	3680	920	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table approximates pass-through of all CWA grants by weighting together pass-through estimates for compliant and non-compliant municipalities. Column 1 estimates $C_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{ts} + \varepsilon_{its}$ on stacks defined by non-compliant municipalities. D_{it} is a dummy variable that equals one after grant receipt, we use D_{it} as a instrument for grant amount. Column 2 repeats this process for compliant municipalities. Column 3 presents a weighted average of these two coefficients where compliant municipalities represent 12 percent of the estimate and non-municipalities municipalities represent the other 88 percent.

Table 1.17: Selection on Gains

	(1) 1972-1974	(2) 1975-1977	(3) 1978-1985	(4) Full Sample	(5) Full Sample
	Sewerage Capital Spending				Grant Amount
Treat * Post	6.36	6.13	4.244	5.86***	12.62***
				(2.15)	(1.779)
Observations	3630	1656	495	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $\theta_S = \sum_{g=1972}^{1985} \tilde{\theta}(g)P(G=g)$ where $\tilde{\theta}(g) = \frac{1}{1985-g+1}\sum_{t=1972}^{1985} 1\{g \leq t\}ATT(g,t)$. Column 1 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1972-1974. Column 2 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1975-1977, and column 3 presents the average of the $\tilde{\theta}(g)$ for municipalities treated in 1978-1981. Column 4 estimates θ_S for the entire sample and column 5 estimates the associated first stage.

Table 1.18: Full Sample Results (Adjusted for Local Matching)

	Grant Amount	Sewerage Capital	
Treat * Post	26.67***	8.84***	
	(3.783)	(2.412)	
Pass-through			0.331***
			(0.106)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 1.2 after adjusting grants for local matching.

Table 1.19: Grant Pass-Through for Sub-Groups (Adjusted for Local Matching))

Panel A: Non-Compliant and < \$125	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	9.31***	7.23	
	(0.980)	(7.701)	
Pass-through			0.782
			(0.853)
p-value: Pass-through = 0			0.359
p-value: Pass-through = 1			0.798
Observations	2162	2162	2162
Panel B:Non-Compliant and >= \$125			
Treat * Post	61.84***	11.81	
	(8.006)	(7.343)	
Pass-through			0.191
			(0.120)
p-value: Pass-through = 0			0.112
p-value: Pass-through = 1			0.000
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	22.20***	0.553	
	(5.619)	(4.801)	
Pass-through			0.025
			(0.286)
p-value: Pass-through = 0			0.930
p-value: Pass-through = 1			0.001
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 1.4 after adjusting grants for local matching.

Table 1.20: Average Pass-Through

	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	21.73***	7.429	
	(3.487)	(5.943)	
Average Pass-through			0.212
			(1.070)
p-value: Pass-through = 0			0.843
p-value: Pass-through = 1			0.462
Observations	4600	4600	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{is} + \varepsilon_{its}$ on a pooled sample of all stacks. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 shows the average pass-through rate obtained from dividing the each stack's reduced form by its first stage.

Table 1.21: Full Sample Results (Never-Treated)

	Grant Amount	Sewerage Capital	
Treat * Post	20.58***	5.274***	
	(1.31)	(1.660)	
Pass-through			0.256***
			(0.079)
p-value: Pass-through = 0			0.001
p-value: Pass-through = 1			0.000
Observations	8492	8492	8492

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 1.2 using never-treated units as a control group

Table 1.22: Grant Pass-Through for Sub-Groups (Never-Treated)

	(1)	(2)	(3)
Panel A: Non-Compliant and < $125	Grant Amount	Sewerage Capital	
Treat * Post	3.53***	3.35***	
	(.230)	(1.087)	
Pass-through			0.974***
			(.310)
p-value: Pass-through = 0			0.002
p-value: Pass-through = 1			0.865
Observations	4494	4494	4494
Panel B:Non-Compliant and >= $125			
Treat * Post	22.72***	8.131***	
	(1.756)	(1.142)	
Pass-through			0.376***
			(.058)
p-value: Pass-through = 0			0.000
p-value: Pass-through = 1			0.000
Observations	3940	3940	3940
Panel C: Compliant			
Treat * Post	11.69***	1.226	
	(1.597)	(1.753)	
Pass-through			0.105
			(0.160)
p-value: Pass-through = 0			0.512
p-value: Pass-through = 1			0.000
Observations	3476	3476	3476

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 1.4 using never-treated units as a control group.

Table 1.23: Observable Characteristics Do Not Predict Timing of Grant Receipt

	(1)	(2)	(3)
Baseline Year:	1971	1974	1977
Compliant Facility	0.408	1.211	-0.641
	(0.777)	(1.804)	(2.137)
ln Population	-0.823**	0.0214	0.400
	(0.342)	(0.395)	(0.971)
Total Revenue	-0.00813	0.0130	0.0347
	(0.0185)	(0.0133)	(0.0356)
Total IGR	-0.00629	0.00447	0.000202
	(0.0105)	(0.0139)	(0.0264)
Federal IGR	0.00482	-0.0141	0.00558
	(0.0118)	(0.0161)	(0.0553)
Nonwater Utility Rev	0.0406	0.0116	-0.00982
	(0.0287)	(0.0249)	(0.0344)
Property Tax Revenue	0.00822	0.0297	-0.0256
	(0.00796)	(0.0227)	(0.0298)
Total Expenditure	0.00872	-0.0116	-0.0486
	(0.0177)	(0.0127)	(0.0494)
Total IG Exp	-0.0186	-0.109*	0.0598
	(0.0303)	(0.0554)	(0.0691)
Capital Outlays	0.00992	0.00255	0.0427
	(0.0145)	(0.0123)	(0.0429)
Highway Exp	-0.0141	0.00760	-0.0404
	(0.0289)	(0.0677)	(0.0723)
Sewerage Capital Outlay	-0.0120	0.0276	0.00423
	(0.0233)	(0.0239)	(0.0464)
Sewerage Non-capital Exp	0.0129	-0.0114	-0.00647
	(0.100)	(0.175)	(0.133)
Nonwater Utiltiy Exp	-0.0359	-0.00549	0.0140
	(0.0264)	(0.0208)	(0.0293)
Water IG Exp	0.159	-0.197	-0.0221
	(0.109)	(0.203)	(0.973)
Water Non-capital Exp	-0.0370	-0.0406	0.241
	(0.0758)	(0.0916)	(0.315)
Water Capital Outlay	-0.0890***	-0.0194	-0.0138
	(0.0337)	(0.0361)	(0.0897)
Water Utility Revenue	-0.0312	0.0476	-0.158
	(0.0382)	(0.0561)	(0.212)
p value	0	0	.08
Observations	112	52	35

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents results from regressing the year a municipality receives its first CWA grant on potential predictive covariates. Column 1 estimates $GrantYear_i = \alpha_0 + compliant_i + X_{ip} + \varepsilon_{ip}$ for all treated municipalities using covariate data from 1971. Columns 2 and 3 use data from 1974 and 1977, respectively, and limit the data to municipalities that do not receive a grant until after those years. All covariates aside from the facility compliance dummy and natural log of population represent per capita values. We present p values from a test of joint significance of all variables at the bottom of the table.

Table 1.24: Full Sample Results (Stacked Difference-in-Differences)

	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	21.73***	7.429	
	(3.487)	(5.943)	
Pass-through			0.342
			(0.267)
p-value: Pass-through = 0			0.200
p-value: Pass-through = 1			0.014
Observations	4600	4600	4600

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $Y_{it} = \alpha_0 + \beta D_{it} + \alpha_{is} + \alpha_{is} + \varepsilon_{its}$ on a pooled sample of all stacks. Grant amount is the dependent variable in column 1, and sewerage capital spending is the dependent variable in columns 2 and 3. Column 3 shows the implied pass-through rate from dividing the first stage in column 1 and the reduced form in column 2.

Table 1.25: Grant Pass-Through for Sub-Groups (Dynamic CS Estimates)

	(1)	(2)	(3)
Panel A: Non-Compliant and $<$ \$125	Grant Amount	Sewerage Capital	
Treat * Post	4.820***	5.973*	
	(0.638)	(3.338)	
Pass-through			1.239
			(0.774)
p-value: Pass-through = 0			.110
p-value: Pass-through = 1			.758
Observations	1410	1410	1410
Panel B: Non-Compliant and $>=$ \$125			
Treat * Post	53.07***	9.25***	
	(13.868)	(3.432)	
Pass-through			0.174*
			(0.092)
p-value: Pass-through = 0			.059
p-value: Pass-through = 1			0.000
Observations	1353	1353	1353
Panel C: Compliant			
Treat * Post	16.17***	7.86	
	(3.160)	(6.196)	
Pass-through			0.486
			(0.415)
p-value: Pass-through = 0			.242
p-value: Pass-through = 1			0.215
Observations	726	726	726

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 1.2 on sub-samples of municipalities. Panel A summarizes pass through for non-compliant municipalities that received grants totaling less than \$125, Panel B for all other non-compliant municipalities, and Panel C for all compliant municipalities.

Table 1.26: Grant Pass-Through for Sub-Groups (Simple CS Estimates)

Panel A: Non-Compliant and < \$125	(1) Grant Amount	(2) Sewerage Capital	(3)
Treat * Post	4.140***	3.139	
	(0.494)	(2.478)	
Pass-through			0.758
			(0.759)
p-value: Pass-through = 0			0.318
p-value: Pass-through = 1			0.750
Observations	1320	1320	1320
Panel B:Non-Compliant and >= \$125			
Treat * Post	52.30***	7.386**	
	(12.281)	(3.031)	
Pass-through			0.141*
			(0.077)
p-value: Pass-through = 0			0.067
p-value: Pass-through = 1			0.000
Observations	888	888	888
Panel C: Compliant			
Treat * Post	12.471***	5.728	
	(2.181)	(4.702)	
Pass-through			0.459
			(0.406)
p-value: Pass-through = 0			0.259
p-value: Pass-through = 1			0.183
Observations	693	693	693

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel D of Table 1.2 on sub-samples of municipalities. Panel A summarizes pass through for non-compliant municipalities that received grants totaling less than \$125, Panel B for all other non-compliant municipalities, and Panel C for all compliant municipalities. We present the p value from testing for a pass-through rate of one at the bottom of the table.

Table 1.27: Pre-CWA Sewerage Capital Spending

	(1) Full Sample	(2) Small Non-compliant	(3) Large Non-compliant	(4) Compliant
Total Sewerage Capital Spending	2,476,679	3,410,038	2,278,557	851,912
Observations	1760	752	656	352

Notes: This table shows average total sewerage capital spending from 1955 to 1971 for the full sample of municipalities, non-compliant municipalities that received grants totaling less than \$125 per capita, non-compliant municipalities that received grants totaling more than \$125 per capita, and compliant municipalities.

Table 1.28: Water Revenue Only Decreased in Compliant Municipalities

	(1) Compliant	(2) Non-Compliant
Panel A: Simple Aggregation	Water Revenue	
Treat * Post	-8.825***	0.653
	(1.131)	(0.953)
Observations	726	2904
Panel B: Dynamic Aggregation		
Treat * Post	-5.873***	0.352
	(1.161)	(0.963)
Observations	726	2904

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents estimates of the effect of grant receipt on water revenue. Panel A presents averages of group-time treatment effects based on group size calculated with the following equation $\frac{1}{\kappa}\sum_{g=1972}^{1985}\sum_{t=1972}^{1987} 1\{g \leq t\}\widehat{ATT(g,t)}P(G=g)$ where $\kappa = \sum_{g=1972}^{1985}\sum_{t=1972}^{1987} 1\{g \leq t\}P(G=g)$. Panel B presents averages of the $\widehat{\theta}_D(e)$ for each time relative to treatment e for $e \geq 0$ calculated with $\widehat{\theta}_D(e) = \sum_{g=1972}^{1985}\sum_{t=1957}^{1987} 1\{t-g=e\}\widehat{ATT(g,t)}P(G=g|t-g=e)$. Column 1 presents results for compliant municipalities with not-yet-treated compliant municipalities acting as controls, and column 2 presents results for non-compliant municipalities with not-yet-treated non-compliant municipalities acting as controls.

Table 1.29: Municipalities Raised Water Bills When Grants Were Too Small

	(1) Less than Costs	(2) Greater than Costs
	Water Revenue	
Treat * Post	1.1798	-0.2012
	(1.7285)	(0.8005)
Observations	960	825

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estimates the results in column 2 of Table 1.28 on sub-samples of municipalities where grants were larger or smaller than estimated costs. Column 1 shows results for non-compliant municipalities where costs were greater than grant amount, and column 2 shows results for non-compliant municipalities where costs were less than grant amount.

Table 1.30: Main Water Revenue Estimates with Alternative Data

	(1) Water Revenue	(2) Water Revenue	(3) Water Revenue
Grant X Compliant	-2.500*	-2.488*	-2.316*
	(1.277)	(1.288)	(1.189)
Timing Group X Year FE	X	X	
Region X Year FE		X	
Observations	6412	6412	6412

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 1.9 on an alternative dataset. Column 1 presents estimates of θ from $R_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Column 2 re-estimates the result in column 1 with region-by-year fixed effects. Column 3 re-estimates the result in column 1 with year fixed effects instead of timing-group-by-year fixed effects.

Table 1.31: Timing Water Revenue Estimates with Alternative Data

	(1) Compliant	(2) Non-Compliant
Panel A: Simple Aggregation	Water Revenue	
Treat * Post	-7.212	0.366
	(5.281)	(2.378)
Observations	1881	6363
Panel B: Dynamic Aggregation		
Treat * Post	-7.122	-1.008
	(13.493)	(2.907)
Observations	1881	6363

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estimates the results in Table 1.28 on an alternative dataset.

Table 1.32: How Much Crowd-Out can Redistribution Account For?

	(1) sewerage capital	(2) water revenue	(3) sewerage capital
grant X compliant	-4.786*	-2.323**	
	(2.450)	(1.157)	
water capital			0.485
			(0.358)
grant X compliant X e	-0.512**	-0.364***	
	(0.245)	(0.138)	
water capital			0.712
			(0.433)
Observations	4752	4752	4752

Standard errors in parentheses, clustered at the municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: Panel A presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i + \alpha_i + \alpha_{gt} + \varepsilon_{it}$. Panel B presents estimates of θ from $Y_{it} = \alpha_0 + \theta grant_{it} * compliant_i * e_{it} + \alpha_i + \alpha_{gt} + \varepsilon_{it}$ where e_{it} indicates year relative to treatment. The dependent variable is sewerage capital spending in column 1 and water revenue in columns 2 and 3. In column 3, we use $grant_{it} * compliant_i$ as an instrument for sewerage capital spending

Table 1.33: Bacon Decomposition of Sewerage Capital Spending TWFE Estimate

	(1) Weight	(2) Average Estimate
Newly Treated vs Not-Yet-Treated	0.454	4.071
Newly Treated vs Already-Treated	0.546	2.512

Notes: This table presents the average difference-in-difference estimates and their weights from decomposing β_{rf}^{TWFE} from $C_{it} = \alpha_0 + \beta_{rf}^{TWFE} D_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ using Goodman-Bacon et al. (2019).

Table 1.34: Dose-Response TWFE Estimates

	(1) Small Non-Compliant	(2) Large Non-Compliant	(3) Compliant
	Cumulative Sewerage Capital Spending		
Cumulative Grant Dollars	1.045**	0.133	0.410*
	(0.517)	(0.100)	(0.214)
p value: Pass-through = 0	0.043	0.184	0.056
p value: Pass-through = 1	0.931	0.000	0.059
Observations	2068	1804	968

Standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of $C_{it} = C_{it} = \alpha_0 + \beta_{dr}^{TWFE} g_{it} + \alpha_i + \alpha_t + \varepsilon_{it}$ on sub-samples of municipalities. C_{it} is per capita sewerage capital expenditure, g_{it} is per capita grant amount, and α_i and α_t are municipality and year fixed effects. Column 1 summarizes pass through for non-compliant municipalities that received grants totaling less than \$125, column 2 for all other non-compliant municipalities, and column 3 for all compliant municipalities.

Table 1.35: Definition of Compliance

	EPA Standard Only	Higher Standard	Full Sample
At or Above Standard	5,872	1,113	6,985
Below Standard	4,613	3,334	7,947
Total	10,485	4,447	14,932

Notes: This table

Table 1.36: Definition of Compliance for Merged Facilities

	EPA Standard Only	Higher Standard	Full Sample
At or Above Standard	1,603	579	2,182
Below Standard	2,424	1,830	4,254
Total	4,027	2,409	6,436

Notes: This table presents the results from Table 1.35 for facilities that received CWA grants.

Table 1.37: Stacked Difference-in-Difference Pass-Through Estimates (EPA Mandate Only)

	(1)	(2)	(3)
Panel A: Non-Compliant and $< \$125$	Grant Amount	Sewerage Capital	
Treat * Post	7.717***	9.308	
	(1.076)	(7.525)	
Pass-through			1.176
			(0.997)
p-value: Pass-through = 0			0.238
p-value: Pass-through = 1			0.865
Observations	1656	1656	1656
Panel B: Non-Compliant and $>= \$125$			
Treat * Post	56.71***	19.95**	
	(11.667)	(7.662)	
Pass-through			0.352**
			(0.145)
p-value: Pass-through = 0			0.015
p-value: Pass-through = 1			0.000
Observations	1150	1150	1150
Panel C: Non-Compliant and $< \$167$			
Treat * Post	8.808***	9.410	
	(1.215)	(7.852)	
Pass-through			1.068
			(0.971)
p-value: Pass-through = 0			0.272
p-value: Pass-through = 1			0.936
Observations	1932	1932	1932
Panel D: Non-Compliant and $>= \$167$			
Treat * Post	61.61***	21.18***	
	(11.851)	(7.71)	
Pass-through			0.344***
			(0.141)
p-value: Pass-through = 0			0.015
p-value: Pass-through = 1			0.000
Observations	874	874	874

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panels B and C of Table 1.4 on the sample of municipalities that did not have to satisfy state wastewater capital requirements greater than the EPA's secondary treatment requirement. Panel A summarizes pass-through for stacks defined by non-compliant municipalities who received grants totalling less than $125, and Panel B summarizes pass-through for all other stacks defined by non-compliant municipalities. Panel C summarizes pass-through for stacks defined by non-compliant municipalities that received grants totalling less than $167, and Panel D repeats this for all other stacks defined by non-compliant municipalities

Table 1.38: Non-Linear Least Squares Estimate of Break in Cost (EPA Mandate Only)

	(1)
	Total Cost
Split	167.3
	(237.7)
Observations	29

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table re-estiamtes the results in Table 1.3 on a sample of non-compliant municipalities that did not face state level treatment technology requirements greater than the EPA's secondary treatment standard.

Table 1.39: Full Sample Results (Nominal)

	Grant Amount	Sewerage Capital	
Treat * Post	26.31^{***}	13.60^{***}	
	(3.589)	(3.200)	
Pass-through			0.517
			(0.126)
p-value: Pass-through = 0			0.000
p-value: Pass-through = 1			0.000
Observations	3630	3630	3630

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Panel B of Table 1.2 measuring spending and grant amount in nominal terms.

Table 1.40: Grant Pass-Through for Sub-Groups (Nominal)

	(1)	(2)	(3)
Panel A: Non-Compliant and $< \$125$	Grant Amount	Sewerage Capital	
Treat * Post	6.72***	9.37	
	(0.747)	(7.740)	
Pass-through			1.39
			(1.21)
p-value: Pass-through = 0			0.251
p-value: Pass-through = 1			0.745
Observations	2162	2162	2162
Panel B:Non-Compliant and $>= \$125$			
Treat * Post	50.75***	18.55**	
	(6.32)	(9.456)	
Pass-through			0.365*
			(0.188)
p-value: Pass-through = 0			0.052
p-value: Pass-through = 1			0.001
Observations	1518	1518	1518
Panel C: Compliant			
Treat * Post	30.39***	14.30	
	(4.622)	(8.10)	
Pass-through			0.471*
			(0.269)
p-value: Pass-through = 0			0.081
p-value: Pass-through = 1			0.049
Observations	920	920	920

Bootstrap standard errors in parentheses, clustered at municipality level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 1.4 measuring dollars in nominal terms.

Table 1.41: Water Revenue Results (Nominal)

	(1)	(2)	(3)	(4)
	Water Revenue	Water Revenue	Water Revenue	Water Revenue
Grant X Compliant	-5.353**			
	(2.587)			
Grant X Compliant X e		-0.850**	-0.956**	-1.233***
		(0.376)	(0.424)	(0.442)
Timing Group X Year FE	X	X	X	
Region X Year FE			X	
Observations	4752	4752	4752	4840

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the results in Table 1.9 measuring water revenues in nominal terms.

CHAPTER 2

A Watershed Moment: The Clean Water Act and Infant Health

The Clean Water Act is a landmark, yet controversial, policy. The CWA caused significant reductions in pollution, but improvements in surface water quality stemming from the CWA have come at a high cost; projects funded through grants to wastewater treatment facilities between 1960 and 2005 cost about $870 billion over their lifetimes (in 2017 dollars) (Keiser and Shapiro, 2019b). In total, US government and industry have spent over $1.9 trillion to abate surface water pollution (Keiser et al., 2019). Existing analyses of the Clean Water Act estimate benefits that are lower than the Act's costs (e.g. Lyon and Farrow (1995); Freeman (2010); Keiser et al. (2019)), but these analyses have not generally included improvements in health caused by the Clean Water Act because there has not been a systematic ex-post measurement of the health benefits of the CWA. To our knowledge, this paper is the first to estimate the effect of CWA grants on infant health and incorporate infant health benefits into a cost-benefit analysis of the CWA.

Incorporating health benefits into a cost-benefit analysis may matter for several reasons. Historically, policies targeting improvements in child health generate high returns to public funds (Hendren and Sprung-Keyser, 2020), and previous economics literature shows that even small increases in child and infant health can lead to large improvements in later life outcomes (Behrman and Rosenzweig, 2004; Royer, 2009; Black et al., 2007; Figlio et al., 2014). Health benefits often account for a large portion of the total benefits of environmental regulation, with health effects accounting for over 95 percent of all benefits of air pollution regulation (Keiser et al., 2019).

Existing economics research estimates the benefits of improved surface water using hedonic analysis that measures the effect of CWA grants on nearby housing prices. Comparing waters up and downstream from wastewater treatment facilities, Keiser and Shapiro (2019a) find that CWA grants caused reductions in downstream pollution. These improvements in water quality were capitalized into housing prices, but increases in home values were substantially smaller than the CWA's costs. By quantifying how residents value water quality, Keiser and Shapiro (2019a) improve upon previous cost-benefit calculations, however, as noted in Keiser et al. (2019), hedonic analysis assumes housing values reflect the implicit value that households place on the quality of nearby surface water. If households are uninformed about nearby surface water quality or do not understand the benefits of reduced surface water pollution, housing values will not reflect the health benefits of the program. In this historic context, it is unlikely that households fully understood the range and extent of the negative health effects of surface water contamination, especially the negative impacts on developing fetuses in utero. By directly estimating the health effects of the CWA, our results complement those in Keiser and Shapiro (2019a) by quantifying one of the largest benefits of the CWA that hedonic analysis is least likely to capture.

Using a difference-in-differences design, we compare infant health outcomes upstream and downstream from wastewater treatment facilities before and after the facility receives a CWA grant. Comparing up and downstream births addresses the endogenous distribution of grants as well as any economic shocks caused by grant receipt, but estimates may still be biased if individuals sort into downstream areas or if these areas experience differential trends relative to upstream areas after grant receipt. To address this concern, we show that CWA grants only caused improvements in surface water quality downstream from facilities that were required to upgrade their treatment technology to comply with new treatment technology standards imposed by the CWA. This finding motivates a triple difference design that uses counties up and downstream from facilities where these treatment technology requirements were not binding as an additional control group. By using already compliant facilities that receive grants as an additional control group, we can account for differential sorting into downstream areas after grant receipt, so the health benefits we capture with this design are likely caused by improvements in water quality.

Across specifications, we consistently find that CWA grants had a statistically significant impact on downstream birth weight. Our results show that reductions in surface water pollution from the CWA are associated with an 8 gram increase in average birth weight. While the monetary benefits of these improvements to infant health are substantial, incorporating infant health alone does not alter the conclusion of a cost-benefit analysis of the CWA. A back-of-the-envelope calculation bounds the monetary benefits of the CWA on infant health under $32 billion, 21 percent of the amount necessary to consider the Clean Water Act grants program cost-effective.

Our results contribute to a literature documenting the importance of effective sewerage and clean water in protecting health historically in the US (Troesken, 2001, 2002; Cutler and Miller, 2005; Ferrie and Troesken, 2008; Beach et al., 2016; Anderson et al., 2020), as well as a literature on the complementarity of sewerage and clean water interventions. Examining water policy in early 20th century Massachusetts, Alsan and Goldin (2019) show that mortality declines were driven by a combination of clean water initiatives and effective sewerage. Watson (2006) shows that federal sanitation policies explain much of the convergence in Native American and White infant mortality rates in the US since 1970. By improving sewerage systems and reducing pollution of surface water throughout the US at a time when most publicly provided drinking water had some treatment, the CWA provides a new context to examine the effect of improved water quality on health in the late 20th century, as well as the complementarity between sewerage infrastructure and clean water in protecting health nationwide.

2.1 Background

The Clean Water Act aimed to slow the flow of contaminants from point sources, such as municipal waste treatment facilities and industrial pollution sources, into rivers and lakes.[1] This paper focuses on CWA grants distributed to municipal wastewater treatment facilities. Wastewater from homes, businesses, and industries, as well as surface runoff, is typically collected through a system of sewers and delivered to a wastewater treatment facility for treatment and discharge into local waterways (USEPA, 2004). The resulting municipal waste is almost entirely organic (Hines, 1966), and often contains both pathogenic and nonpathogenic microorganisms harmful to human health.[2] Human exposure to contaminated surface water could occur through several pathways, including direct contact during recreational activities, through consuming or bathing with contaminated private well water, or through consuming or bathing with contaminated public drinking water, especially when public water is obtained from surface water sources.[3] To reduce surface water contamination, the CWA addressed pollution from municipal waste treatment plants with two complementary policies: grants to wastewater treatment facilities, and regulation of wastewater treatment technology.[4]

2.1.1 Grants

From 1972 to 1988, the EPA distributed an estimated $153 billion (in 2014 dollars) in grants to municipal governments for capital upgrades to wastewater treatment facilities. The EPA allocated CWA grant money to states according to a formula based on total population, forecast population, and wastewater treatment needs (Rubin, 1985). States then distributed grants to municipalities according to priority lists based on the severity of nearby surface water pollution, the size of the population affected, the need for conservation of the affected waterway, and that waterway's specific category of need (USEPA, 1980).

Since state governments wrote their own priority lists, grant placement may be correlated with trends in infant health. Moreover, grants could cause increases in birth weight that are unrelated to changes in pollution by improving economic conditions with an influx of federal dollars. Instead of treating grant timing and location as exogenous, we compare the difference in birth outcomes in areas up and downstream from a given wastewater treatment facility before and after grant receipt between facilities that were required to make treatment technology upgrades and all other facilities. To the extent that other policies were changing during this time period, and that grants improved local economic conditions, these changes were likely to affect upstream and downstream areas similarly.

[1] Although much of the contamination of US waterways comes from sources that cannot be traced back to a specific facility, such as agricultural runoff, the Clean Water Act did not directly regulate these "non-point" pollution sources. The CWA did not directly regulate drinking water supplies either. The Safe Drinking Water Act sets minimum standards for drinking water quality for all public water systems in the US.

[2] Pathogens harmful to human health include enteric bacteria, viruses, protozoa, parastic worms, and their eggs. These microorganisms can cause a range of gastrointestinal illnesses and infections (Reynolds et al., 2008; Chahal et al., 2016).

[3] While data limitations prevent us from definitively determining the main pathway of exposure, we explore potential pathways in section 2.6.2 of the appendix. We find suggestive evidence that public drinking water sourced from surface waters is one channel through which infant health improves. This is consistent with evidence that violations of modern coliform drinking water standards harm health (Marcus, 2021).

[4] In addition to regulating municipal waste treatment facilities, the CWA required all industrial polluters to obtain a permit from the National Pollutant Discharge Elimination System (NPDES) before discharging wastewater. Regulation through the NPDES led to reductions in both profits (Rassier and Earnhart, 2010) and the number of environmental employees (Raff and Earnhart, 2019) at newly regulated polluters.

2.1.2 Regulation

In 1972, about a quarter of all US municipal wastewater treatment facilities reported using relatively inexpensive, but less effective, primary treatment (USEPA, 2000). This process, depicted in Figure 2.4, forces wastewater through a series of screens. While primary treatment removes large detritus and heavy biosolids, it still discharges all but the heaviest organic material into waterways (USEPA, 1998).

The Clean Water Act required all municipal treatment plants to upgrade to secondary treatment. Plants use secondary treatment technology, shown in Figure 2.5, in addition to primary treatment. After screens filter out large debris, wastewater sits in an aeration tank where bacteria in the water consume organic material, which ultimately reduces biochemical oxygen demand (BOD).[5] Through this biological oxidation, secondary treatment can remove over 90 percent of harmful pathogenic bacteria and viruses from sewage (Abdel-Raouf et al., 2012).[6] While it is not practical to monitor pathogens directly, regulators and researchers often use indicator organisms, such as total or fecal coliforms, to monitor water quality. Keiser and Shapiro (2019a) show that grants to wastewater treatment facilities improved key indicators of water quality, including dissolved oxygen deficit, BOD, and fecal coliforms. As dissolved oxygen deficit is the most consistently and widely monitored measure of water quality in our sample, we focus on this measure.

The potential benefits of upgrading a facility's treatment technology were well understood, but waste treatment capital upgrades were expensive. Upgrading to secondary treatment technology could increase a facility's operating costs by up to 60 percent and require capital investments of as much as 30 percent of the initial cost of the facility (National Environmental Research Center, 1972). Because of these costs, 53 percent of plants in the 1972 Clean Watershed Needs Survey (CWNS) were not in compliance with both state and federal treatment technology mandates.[7] Treatment plants that were not in compliance with both state and federal capital mandates in 1972, which we refer to as "non-compliant" facilities, had a strong incentive to use CWA grants to offset the costs of upgrading their treatment technology.[8]

Many facilities that were already in compliance with both state and federal mandates still received CWA grants. While these facilities could make capital improvements, such as increasing capacity, they had less incentive to do so. Since the CWA did not mandate these upgrades, there was no binding constraint requiring these facilities to spend grant money on sewerage capital upgrades, and the municipalities that operated them faced pressure to use grant money to offset the operating costs of their water and sewerage utilities in an attempt to lower costs for consumers and become more competitive (Daigger, 1998).[9]

Since non-compliant facilities had a clear channel through which to improve surface water quality and were more likely to spend CWA grant money on capital upgrades, we expect the reductions in downstream pollution associated with CWA grants to be largest for non-compliant facilities. This motivates a triple difference design that uses areas near facilities that were not indicated as pre-CWA non-compliant in the 1972 CWNS as an additional control group.

2.2 Data

CWA Grants and Municipal Wastewater Treatment Plants

We obtain data on all 33,429 grants that the EPA distributed to 14,285 wastewater treatment plants from the EPA's Grant Information Control System.[10] Most facilities received multiple grants, so we define a facility as "treated" after it receives its first CWA grant.

Using a unique facility code, we merge this grant data with the Clean Watershed Needs Survey, an assessment of the capital investment needed to meet the water quality goals of the CWA. This linked data includes

[5] Additionally, some states required facilities to meet more stringent treatment technology requirements than the CWA's mandate, such as tertiary treatment, which is aimed at removing ammonium, nitrate, and phosphate (USEPA, 2000).

[6] Suspended (e.g. activated sludge) growth reactors remove about 90 percent of viruses, but removal can be more varied in film reactors, which provide less absorption (Abdel-Raouf et al., 2012).

[7] See Appendix Section 2.6.5.4 for more discussion.

[8] Permits distributed to polluters through the NPDES required municipal treatment plants to satisfy the treatment technology mandate, and violators could be fined up to $25,000 per day (Copeland, 2016).

[9] Flynn and Smith (2021) show that CWA grants to non-compliant municipalities led to a dollar for dollar increase in sewerage capital spending, while grants to facilities already in compliance with state and federal capital mandates crowded out funds that municipalities were already spending on sewerage capital rather than causing an increase in sewerage capital spending.

[10] The 33,429 grants in our sample exclude grants that do not include a specific facility code, as it is unclear to what extent these grants were precisely for wastewater treatment plants. Appendix Section 2.6.5.4 provides further discussion.

facility location, grant timing, and state and federal treatment technology compliance status as of 1972.[11]

Spatial Data on Waterways

We define treatment in terms of the flow direction of waterways. We determine if a waterway is up or downstream from a facility with the National Hydrography Data Set, an electronic atlas that maps the location and flow direction of all US waterways. We follow both the EPA and other researchers studying the Clean Water Act by focusing on areas 25 miles up and downstream from treatment facilities (Keiser and Shapiro, 2019a; USEPA, 2001).[12]

Water Pollution

Data on dissolved oxygen deficit comes from STORET legacy, which includes readings from pollution monitoring stations across the US.[13] We include readings from pollution monitors on rivers and lakes located 25 miles up or downstream from any facility in the CWNS data.[14]

Infant Health

We measure infant health with birth certificate data from the National Center for Health Statistics (NCHS) from 1968 to 1988. These data contain information on birth weight, birth order, mother's age and race, and county of residence for each birth.[15] Table 2.4 presents summary statistics for individual-level births in 1970, two years before the first CWA grants were distributed, from up and downstream counties.

While ideal data would contain exact addresses, these data are unavailable for most states, and even when collected, addresses are typically not available until the after the adoption of the 1989 US Standard Birth Certificate revision and the use of electronic birth certificates, which is after our study period.

We collapse birth data to county means, calculating the average birth weight, the probability of low birth weight, the percent of non-white births, average mother's age, and the probability of being a mother's first, second, third, or fourth or higher birth in each county year.[16] We define a county as downstream if it contains any waterway that is within 25 miles downstream of a treated facility. Although more recent birth records data contains far more variables of interest, such as gestation, maternal education, and maternal risk factors, our analysis relies on data from 1968 to 1988 and these variables are either unavailable or not reliably and consistently recorded in data from this time period.

Population Density

We expect the health effects of improved surface water quality to be concentrated near treated waterways. County-level exposure depends on the distribution of the population within a county relative to the location of treated waterways. We use 1990 census block population density data from the US Census Bureau to measure the percent of a county's population living within a mile of a treated waterway.[17] Assuming a uniform population distribution within census blocks, this provides a proxy for the probability that mothers within the county are exposed to treated waterways.

---

[11] There are 1,930 facilities in our analysis sample that are missing data on pre-CWA treatment technology. We assume that these facilities were already in compliance with state and federal treatment technology requirements. Throughout the paper, we refer to the set of "compliant" facilities, which includes all facilities that were not explicitly "non-compliant" in the 1972 CWNS. Our results are similar when we exclude facilities with missing information on treatment technology.

[12] Table 2.13 shows that our results are robust to concentrating on areas 5 or 10 miles downstream from treatment facilities.

[13] Dissolved oxygen deficit is a continuous measure defined as 100 minus dissolved oxygen saturation (dissolved oxygen level divided by water's maximum oxygen level). It is one of the most common measures of omnibus water pollution in research, and it responds to a wide variety of pollutants (Keiser and Shapiro, 2019a).

[14] We follow the data cleaning steps laid out in the appendix of Keiser and Shapiro (2019a).

[15] Data before 1972 constitutes a 50 percent random sample of all births in the US. After 1972, some states report data on all births. Six states had full sample data in 1972, and all states and the District of Columbia had full sample data by 1985. Appendix 2.6.5 provides additional information and shows our main results are not driven by sampling changes.

[16] We also calculate county means of one year mortality using data from NCHS (National Center for Health Statistics, 1988b). We find no significant effect of CWA grants on this outcome in Table 2.12, however our estimates are imprecise.

[17] We use data from 1990, because it is the first census for which population density data is available at the census block level. While it would be preferable to know the population density of mothers specifically, this population distribution is likely highly correlated. Appendix 2.6.4.2 shows that our results are robust to scaling by the percent of a county's population living within other bandwidths around treated waterways. We also show similar but attenuated results if we define treatment with a binary variable in Appendix 2.6.4.4.

2.3 First Stage: Water Pollution

2.3.1 Methods

Before comparing birth outcomes up and downstream from wastewater treatment facilities, we examine the first stage relationship between grant receipt and downstream water quality with equation 2.1.

$$Q_{pdy} = \gamma g_{py} * d_d + \beta W_{pdy} + \alpha_{py} + \alpha_{pd} + \varepsilon_{pdy} \tag{2.1}$$

Q_{pdy} is a measure of dissolved oxygen deficit and g_{py} equals one after a facility receives its first CWA grant. There are two observations for each treatment plant p for each year y, which describe average dissolved oxygen deficit upstream ($d_d = 0$) and downstream ($d_d = 1$) from that plant. Since dissolved oxygen deficit varies inversely with temperature, W_{pdy} measures water temperature.

We include plant-by-downstream and plant-by-year fixed effects, α_{pd} and α_{py}, respectively. Plant-by-downstream fixed effects allow waters up and downstream from a given wastewater treatment plant to have different mean levels of dissolved oxygen deficit, which controls for pollution sources located up or downstream from a plant that are constant over time. Plant-by-year fixed effects ensure that we are only comparing waters up and downstream from the same facility, which controls for any yearly shocks that affect waters both up and downstream from a facility. All standard errors in our pollution estimates are clustered at the facility level.

We estimate equation 2.1 for the full sample and subsamples of compliant and non-compliant facilites, as well as a fully interacted triple difference specification. These estimates give us a sense of how grants and regulations worked together by testing whether pollution evolved differently in waters downstream from non-compliant facilities and compliant facilities after grant receipt.

Table 2.1: Effects on Surface Water Pollution

	(1) Full Sample	(2) Non-Compliant	(3) Compliant	(4) DDD
Grant X Downstream	-0.974***	-1.566***	-0.371	-0.371
	(0.199)	(0.285)	(0.276)	(0.276)
Grant X Downstream X Non-Compliant				-1.196***
				(0.396)
Weather Controls	X	X	X	X
Facility X Downstream Fixed Effects	X	X	X	X
Facility X Year Fixed Effects	X	X	X	X
Observations	114148	46968	67180	114148

Standard errors in parentheses, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table describes the effects of Clean Water Act grants on downstream pollution. Columns 1, 2 and 3 estimate equation 2.1 for areas up and downstream from all facilities in our sample, non-compliant facilities, and all other facilities respectively. Column 4 presents estimates from the associated triple difference: $Q_{pdy} = \gamma_0^{DD} g_y * d_d + \gamma^{DDD} g_y * d_d * t_p + \beta W_{pdy} + \phi W_{pdy} * t_p + \alpha_{py} + \alpha_{pd} + \varepsilon_{pdy}$ where t_p is a dummy variable equaling one for observations from non-compliant facilities. Q_{pdy} is dissolved oxygen deficit, g_y is a dummy variable equaling one after a facility receives a CWA grant, and d_d is a dummy equaling one for observations downstream from a facility. All regressions include controls for water temperature, as well as facility-by-downstream and facility-by-year fixed effects, α_{pd} and α_{py}.
Source: (USEPA, 1988)

2.3.2 Pollution Results

Table 2.1 estimates the effect of CWA grant receipt on downstream water quality. Columns 1-3 present estimates of equation 2.1 on the full sample, non-compliant facilities, and compliant facilities, respectively.

Column 4 presents coefficients from a triple difference specification. As shown in column 2, dissolved oxygen deficit only decreased significantly in water downstream from non-compliant facilities. Since dissolved oxygen deficit is defined as 100 minus dissolved oxygen saturation, this result show that, after grant receipt, dissolved oxygen saturation increased by 1.6 percentage point in waters downstream from non-compliant facilities relative to waters upstream from the same facility. The coefficient for waters downstream from compliant facilities in column 3 is small and statistically insignificant. The reduction in dissolved oxygen deficit downstream from non-compliant facilities is statistically larger than the change downstream from compliant facilities, as shown by the significant negative triple difference coefficient in column 4.

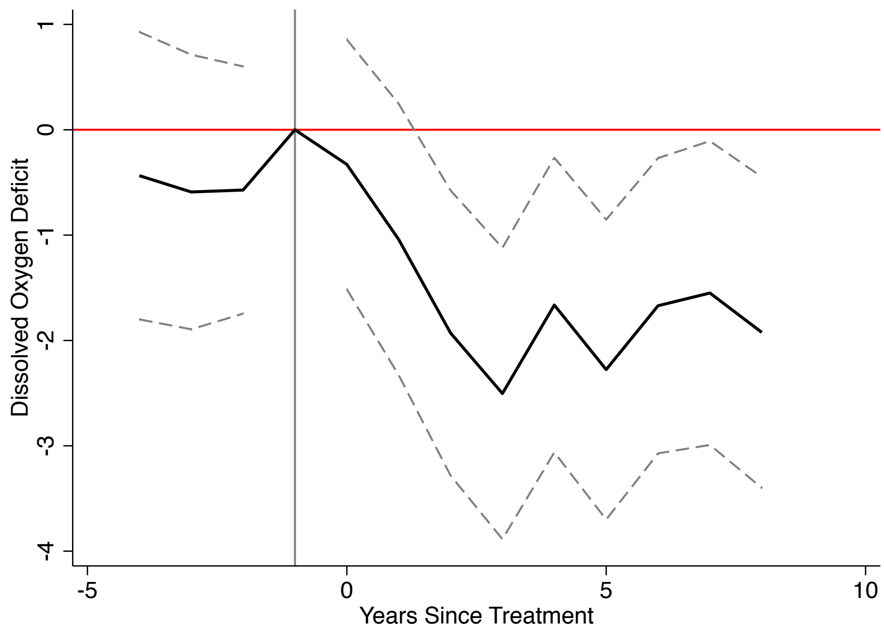


Figure 2.1: Pollution Decreased Downstream from Non-Compliant Grant Facilities

Notes: This figure plots the θ_t and η_t from estimating $Q_{pdy} = \sum_{t=-5}^{-2} \theta_t 1\{y - y_p^* = t\} * d_d * t_p + \sum_{t=0}^{9} \eta_t 1\{y - y_p^* = t\} * d_d * t_p + \sum_{t=-5}^{-2} \pi_t 1\{y - y_p^* = t\} * d_d + \sum_{t=0}^{9} \gamma_t 1\{y - y_p^* = t\} * d_d + \beta W_{pdy} + \phi W_{pdy} * t_p + \alpha_{py} + \alpha_{pd} + \varepsilon_{pdy}$. Q_{pdy} measures dissolved oxygen deficit, d_d is a dummy equaling one for observations downstream from a facility, and t_p is an indicator that equals one for non-compliant facilities. The model includes facility-by-downstream fixed effects and facility-by-year fixed effects, α_{pd} and α_{py}, as well as controls for temperature.
Source: USEPA (1988)

Figure 2.1 presents results from the event study corresponding to the triple difference in column 4. This figure shows that reductions in downstream pollution were significantly larger in waters downstream from non-compliant facilities relative to waters downstream from compliant facilities. In addition, there does not appear to be any trend in pollution prior to grant receipt, which might have arisen from compliant facilities early adoption of more advanced treatment technology. In the analysis of the impact of CWA grants on infant health that follows, we leverage this comparison between non-compliant and compliant facilities in a triple difference specification.

2.4 Infant Health

2.4.1 Methods

We begin our reduced-form analysis of the impact of CWA grants on infant health by comparing birth outcomes in counties downstream from treated facilities to all other areas with the following difference-in-differences specification

$$Y_{cy} = \gamma pct_{cy} + \beta X_{cy} + \alpha_c + \alpha_y + \varepsilon_{cy} \tag{2.2}$$

Y_{cy} is an average birth outcome in county c in year y, and pct_{cy} is the percent of county c's population living within a mile of a treated waterway in year y. Controls in X_{cy} include the percent of births that were a mother's first, second, third, or fourth, and county averages of mother's age and race. α_c and α_y are county and year fixed effects. Observations are at the county-year level and standard errors are clustered at the county level. Since we collapse birth weight data to county means, we weight our results by the total number of births that occurred in a county-year.

The presence of local area trends specific to a facility's location could mean that an upstream county is only a good counterfactual for a county located downstream from the same facility. We address this concern in our next specification by collapsing our data to the facility rather than the county level. Our outcome variable is now ΔY_{py}, which is equal to the mean birth weight in all counties downstream from a facility minus the mean birth weight in all counties upstream from the same facility in each year. We then estimate the following specification

$$\Delta Y_{py} = \gamma pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py} \tag{2.3}$$

where p indexes facilities, and pct_{py} measures the percent of downstream counties' populations living within a mile of a treated waterway. We include facility and year fixed effects, α_p and α_y, respectively.[18] Standard errors are clustered at the facility level.

This specification requires us to assume that, in the absence of grant receipt, birth outcomes would have evolved similarly in areas up and downstream from the same facility after grant receipt. This assumption would be violated if, for example, downstream areas were experiencing differential sorting patterns or greater economic growth relative to upstream areas, even in the absence of CWA grants. For example, Keiser and

[18] Controls in facility-level specifications are averages from all births in up and downstream counties. Our results are robust to controlling for the difference between average demographic characteristics in up and downstream counties instead.

Shapiro (2019a) show that downstream housing prices increase after grant receipt, which may cause healthier mothers to sort into downstream communities.

Table 2.2: Effects on Demographic Changes

	Non-Compliant	Compliant	DDD
	(1)	(2)	(3)
Panel A.	Percent Non-White		
Pct Pop 1 Mile	-0.0223***	-0.0176***	-0.0176***
	(0.00295)	(0.00270)	(0.00270)
Pct Pop 1 Mile X Non-Compliant			-0.00471
			(0.00400)
Mean	.116	.105	.11
Panel B.	Mother's Age		
Pct Pop 1 Mile	0.126***	0.0784**	0.0784**
	(0.0360)	(0.0324)	(0.0324)
Pct Pop 1 Mile X Non-Compliant			0.0479
			(0.0484)
Mean	24.563	24.569	24.566
Panel C.	Probability First or Second Birth		
Pct Pop 1 Mile	-0.00210	0.00109	0.00109
	(0.00360)	(0.00254)	(0.00254)
Pct Pop 1 Mile X Non-Compliant			-0.00319
			(0.00441)
Mean	.653	.645	.648
Panel D.	Probability Third or Higher birth		
Pct Pop 1 Mile	-0.0105***	-0.00618***	-0.00618***
	(0.00204)	(0.00177)	(0.00177)
Pct Pop 1 Mile X Non-Compliant			-0.00429
			(0.00270)
Mean	.338	.347	.343
Unit and Year Fixed Effects	X	X	X
Collapsed to Facility Level	X	X	X
Observations	34188	48132	82320

Standard errors in pare thesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: Columns 1 and 2 present results from estimating $\Delta x_{py} = \gamma pct_{py} + \alpha_p + \alpha_y + \varepsilon_{py}$ on subsamples of Non-Compliant and Compliant facilities. Δx_{py} is a measure of the difference between demographic characteristic in counties up and downstream from facility p in year y, and pct_{py} is a continuous variable that takes values from zero to one, and indicates the Percent of downstream counties' populations living within a mile of a treated waterway in year y. The model includes facility and year fixed effects, α_p and α_y. Column 3 presents estimates of the associated triple difference, $\Delta x_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py}$, where t_p is an indicator that equals one for Non-Compliant facilities. Each panel represents a different demographic variable. Means of each variable in 1970 from up and downstream counties are reported at the bottom of each panel.
Source: National Center for Health Statistics (1988a)

Table 2.3: Effects on Health at Birth

	Full Sample (1)	Full Sample (2)	Full Sample (3)	Full Sample (4)	Non-Compliant (5)	Compliant (6)	DDD (7)
Panel A	County Average Birth Weight						
Pct Pop 1 Mile	12.80***	6.718***	7.134***	8.999***	13.36***	5.153**	5.153**
	(3.105)	(2.389)	(2.392)	(1.672)	(2.729)	(2.052)	(2.052)
Pct Pop 1 Mile X Non-Compliant							8.211**
							(3.413)
Panel B	Probability Birth Weight $<$ 2500 Grams						
Pct Pop 1 Mile	-0.00288***	-0.000874*	-0.000963*	-0.00177***	-0.00216***	-0.00138**	-0.00138**
	(0.000670)	(0.000523)	(0.000521)	(0.000401)	(0.000602)	(0.000538)	(0.000538)
Pct Pop 1 Mile X Non-Compliant							-0.000780
							(0.000808)
Unit and Year Fixed Effects	X	X	X	X	X	X	X
Demographic Controls		X	X	X	X	X	X
Up/Downstream Counties Only			X	X	X	X	X
Collapsed to County Level	X	X	X				
Collapsed to Facility Level				X	X	X	X
Observations	64239	64239	64008	82320	34188	48132	82320

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table describes the effects of Clean Water Act grants on downstream infant health. In Panel A, the dependent variable is the average birth weight in a county-year, and in Panel B, it is the probability of being born weighing less than 2500 grams. Columns 1-3 present estimates from equation 2.2. Pct Pop 1 Mile is a continuous variable that takes values from zero to one, and indicates the proportion of the population that lived within a mile of a treated waterway in each year. All estimates include unit and year fixed effects. Demographic controls include the percent of a county's births in a given birth order bin, and county averages of mother's age and race and child gender. Columns 1 and 2 use data from every county in the US, while columns 3-7 restrict the sample to counties that are up or downstream from a wastewater treatment facility. Columns 5 and 6 restrict the sample to non-compliant facilities (those that were required to make treatment technology upgrades) and compliant facilities (those that were not), respectively. In columns 1-3, data is collapsed to the county level. In columns 4-7, data is collapsed to the facility level. Columns 4-6 estimate equation 2.3. Column 7 estimates the associated triple difference from equation 2.4. All regressions are weighted by number of births.
Source: National Center for Health Statistics (1988a)

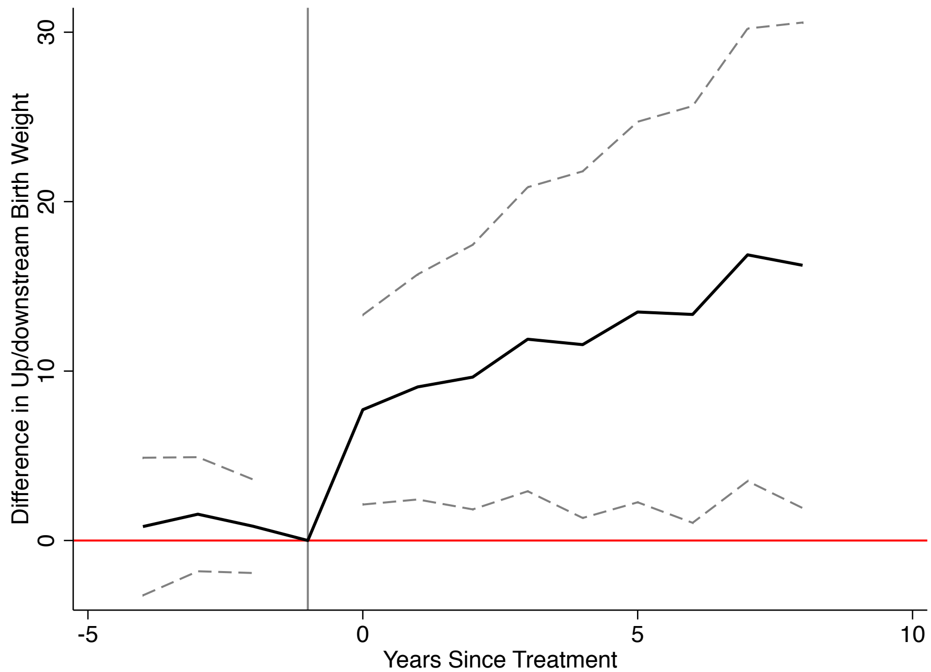


Figure 2.2: Birth Weight Increased Downstream from Non-Compliant Grant Facilities

Notes: This Figure plots the θ_t and η_t from estimating $\Delta Y_{py} = \sum_{t=-5}^{-2} \theta_t 1\{y - y_p^* = t\} * t_p + \sum_{t=0}^{9} \eta_t 1\{y - y_p^* = t\} * pct_{py} * t_p + \sum_{t=-4}^{-2} \pi_t 1\{y - y_p^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_p^* = t\} * pct_{py} + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py}$. Source: National Center for Health Statistics (1988a)

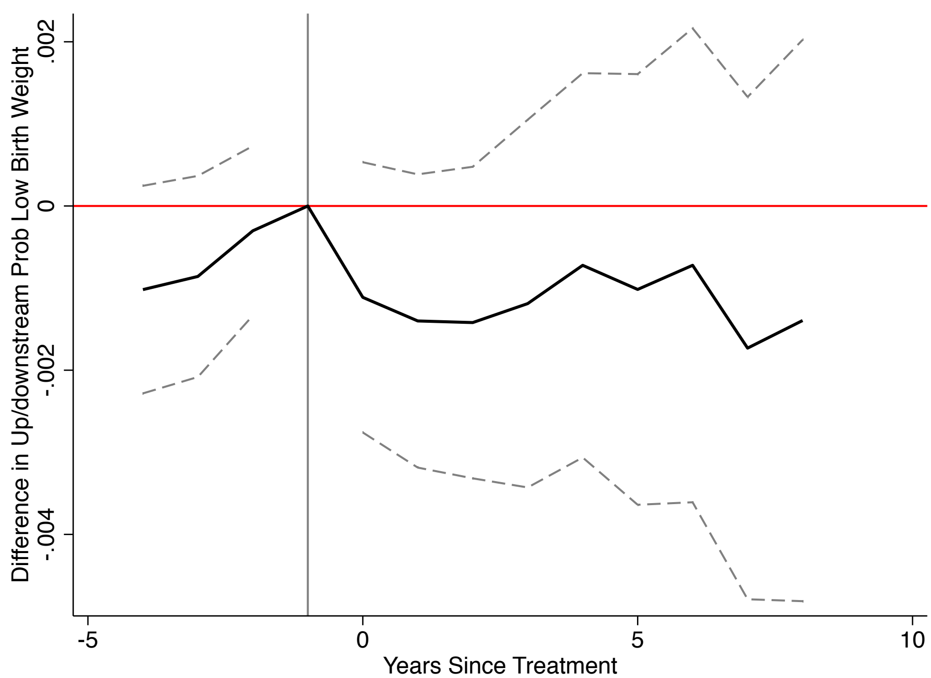


Figure 2.3: Probability of Low Birth Weight Decreased Downstream from Non-Compliant Grant Facilities

Notes: This Figure re-created Figure 2.2 with the difference in the probability of being born weighing less than 2500 grams between up and downstream counties in year y as the dependent variable. Source: National Center for Health Statistics (1988a)

To address concerns regarding differential trends in infant health in downstream relative to upstream areas caused by differences in economic growth or positive sorting of households into downstream areas, we employ a triple difference design. We estimate the following equation

$$\Delta Y_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py} \tag{2.4}$$

where t_p is an indicator that equals one for non-compliant facilities. In this specification, the first difference comes from where and when CWA grants were distributed, the second comes from if a birth occurred up or downstream from a wastewater treatment facility, and the third comes from the facility's compliance with the treatment technology mandate.

Even if individuals sort into downstream communities, so long as the sorting pattern induced by grant receipt is similar for both compliant and non-compliant facilities, using compliance as a third difference will capture unobserved changes to up and downstream counties occurring contemporaneously with CWA grant receipt. We test this by exploring how maternal characteristics evolve after grant receipt in upstream and downstream areas across non-compliant and compliant facilities.

Table 2.2 estimates equation 2.2 on demographic characteristics that are correlated with birth weight, such as race, age, and birth order. Column 1 of Table 2.2 shows results for the subsample of non-compliant facilities and column 2 shows results for compliant facilities. Column 3 presents results from the associated triple difference. Columns 1 and 2 show that areas downstream from facilities that received CWA grants had smaller non-white populations, slightly older mothers, and fewer higher order births, but changes in downstream demographic characteristics are very similar across non-compliant and compliant facilities. The triple difference coefficients presented in column 3 are small and statistically insignificant for all observed demographic outcomes, indicating that there was no observable differential sorting into downstream areas across non-compliant and compliant facilities after grant receipt. These results provide some evidence that the identification assumption for the triple difference specification is likely to hold.

2.4.2 Infant Health Results

Panel A of Table 2.3 shows effects on birth weight that are robust across a variety of specifications. Column 1 compares births in counties downstream from grant facilities to those in any other county by estimating equation 2.2 using a sample of births from every county in the contiguous US. Column 2 adds demographic controls to this specification.[19] Since births occurring in counties that are not near wastewater treatment facilities might not make a good control group, column 3 excludes counties that are not up or downstream

[19]Figures 2.6 and 2.7 in the appendix shows the event study figures the correspond to the estimate in column 2.

from a wastewater treatment facility. This compares births in a downstream county to those in any upstream county. The results are similar to those from the full sample.

Counties upstream from the same facility are likely to make better counterfactuals for downstream counties than counties upstream from any facility. Column 4 estimates equation 2.3, which compares birth weight in counties up and downstream from the same facility. The point estimate is slightly larger in magnitude with a smaller confidence interval.[20,21]

The impact of the CWA on birth weight may not be uniform across the distribution of birth weight, so we also show results for low birth weight in Panel B of Table 2.3. The point estimates are consistently negative, although not always significant, and range from -0.09 to -0.29 percentage points. About 7 percent of births in our sample are low birth weight, so this represents a change of 1 to 4 percent from the mean.

Finally, we estimate our triple difference specification on birth outcomes. Columns 5 and 6 of Table 2.3 present results from estimating equation 2.3 on sub-samples of non-compliant and compliant facilities, respectively. Consistent with our pollution results in Table 2.1, we see a relatively large and statistically significant improvement in birth weight downstream from non-compliant facilities. The effect in areas downstream from compliant facilities is also positive, but smaller; improvements in infant health in areas downstream from compliant facilities may be driven by demographic or economic changes that coincide with grant timing. Since, as shown in Table 2.2, demographic changes were similar in areas downstream from non-compliant and complaint facilities, the difference between the effects downstream from non-compliant and compliant facilities likely comes from the differences in surface water quality shown in Table 2.1, rather than shifting demographics.

Figures 2.2 and 2.3 present the event studies that correspond to our triple difference specification. There is no evidence of pre-treatment trends in infant health outcomes. For birth weight, there is a statistically significant increase in downstream (relative to upstream) counties after a non-compliant facility receives a grant (relative to other facilities).[22] For low birth weight, the point estimates are similar in shape but are less precise. Importantly, the shapes of these event studies are similar to the patterns in pollution shown in Figure 2.1.

[20]Figures 2.8 and 2.9 shows the associated event study figures for column 4.

[21]These results are identified off of comparisons of newly treated facilities relative to never-treated facilities, newly treated facilities relative to facilities that have not yet been treated, and newly treated facilities relative to already-treated facilities (Goodman-Bacon, 2021a). The third type of comparison can be wrong signed. We show in the Appendix sections 2.6.4.3 and 2.6.4.4 that our results are robust to using a stacked difference-in-difference design and Callaway and Sant'Anna (2020a), which only rely on the first two types of comparisons.

[22]In all of our event studies, we report coefficients for four years before and eight years after grant receipt, so that we only report balanced coefficients in our infant health specifications. These specifications also includes bins for five or more years before the grant and nine or more years after the grant, but our results are not sensitive to this choice of binning. While unbalanced event study coefficients should be interpreted with caution, Figure 2.12 presents a version of Figure 2.2 with 16 years of post-treatment data. This figure suggests that the effect of CWA grants on infant health flattens out by 10 years after treatment, consistent with grant projects taking up to 10 years to complete (USEPA, 2002).

We summarize the effect of changes in surface water quality downstream from non-compliant facilities on infant health by estimating equation 2.4 on the pooled sample, which leverages all of our variation in one regression. Since equation 2.4 includes a full set of interactions, our estimate of γ^{DDD}, reported in column 7 of Table 2.3, is equivalent to the difference in the estimates in columns 5 and 6. As in our pollution estimates, the improvements in birth outcomes downstream from non-compliant facilities are statistically larger than improvements downstream from compliant facilities.[23]

The results from this triple difference show that increasing the probability of exposure to treated surface water from zero to 100 percent is associated with an 8.21 gram increase in average birth weight in counties downstream from facilities that were required to make upgrades to their treatment technology. In terms of magnitude, the effect on birth weight is about half of the estimated effect of any exposure to Ramadan during pregnancy (Almond and Mazumder, 2011), and about the same magnitude as the effect of stress in utero due to nearby landmine explosions on birth weight (Camacho, 2008). Estimates of the effect on the probability of low birth weight shown in Panel B of Table 2.3 are not significant, but they do bound improvements above a 0.236 percentage point decrease, or about 3 percent from the mean of low birth weight. This is slightly smaller than the estimated effect of drinking water contamination in utero on low birth weight estimated in a modern context (Currie et al., 2013).

2.5 Discussion & Conclusion

The preceding evidence suggests that the Clean Water Act led to small but significant improvements in infant health, with reductions in pollution associated with CWA grants leading to an 8 gram increase in average birth weight in counties downstream from facilities that were required to make treatment technology upgrades. These results are consistent with the significantly larger improvements in water quality we find downstream from these facilities.

We use this information to incorporate one measure of health benefits into a cost-benefit analysis of the Clean Water Act. In total, CWA grants to wastewater treatment facilities cost an estimated \$153 billion (in 2014 dollars). About 56.4 million births occurred in treated counties between 1972 and 1988, and we estimate that about 32.1 million of those births occurred within a mile of a treated waterway. While our preferred triple difference specification does not show statistically significant changes to the probability of low birth weight, it bounds improvements below a 0.263 percentage point reduction (Column 7 of Table 2.3).

Almond et al. (2005) estimates that low birth weight increases hospital costs by \$8319 and increases 1 year mortality by 37 per 1000 births, and Oreopoulos et al. (2008) finds that low birth weight reduces lifetime

[23]We show that this heterogeneity in effects is not driven by differences in facility size, population served or non-treatment technology upgrades in Table 2.17, which provides further evidence that improvements in downstream infant health are driven by upgrades to treatment technology. In appendix section 2.6.1, we also explore heterogeneity of the main results by maternal race and grant timing, but find no significant differences along these dimensions.

earnings by 3.8 percent. We combine these estimates with the EPA's value of a statistical life (VSL) of \$7.4 million and the census bureau's work-life earnings estimate of \$2.4 million to calculate a back-of-the-envelope estimate of the infant health benefits of the CWA. While a more comprehensive calculation of the health benefits of the CWA would include other potentially impacted health outcomes, we estimate the infant health benefits of the CWA are bounded below \$32 billion, about 21 percent of the amount needed to make the CWA cost effective.[24]

The \$153 billion figure includes grants to compliant facilities, which did not lead to measurable improvements in downstream water quality. If CWA grants had been targeted only towards facilities requiring treatment technology upgrades, the cost-benefit ratio may have been more favorable, as health improvements were detected only downstream of these facilities. Health effects alone can account for as much as 32 percent of the \$101 billion (in 2014 dollars) in grants distributed to non-compliant facilities.

Using increased housing prices to value the benefits of the CWA, Keiser and Shapiro (2019a) estimate a benefit to cost ratio of 0.26. If we assume that hedonic estimates do not capture any health benefits, grants to non-compliant facilities might have a benefit to cost ratio as high as 0.58 after incorporating improvements to infant health. Considering that infant health is only one dimension of health potentially impacted by the Clean Water Act, this is a sizable improvement in the benefit-cost ratio. Including additional measures of health, such as reduced hospital admissions, reduced school absences, and health effects for adolescents and adults would likely increase this ratio even further. More generally, this research documents the importance of policies targeting cleaner water through sewerage treatment and that the complementarity between clean drinking water and sewerage initiatives for improving human health holds well into the twentieth century.

[24]Estimates of VSL vary. Using a VSL of \$10 million from Kniesner and Viscusi (2019) instead of the EPA estimate does not change the conclusion of a cost-benefit analysis.

2.6 Appendix

2.6.1 Heterogeneity

We examine the heterogeneity of our estimates across race in Table 2.5 by estimating equation 2.4 on subsamples of white and non-white births from counties with sizable nonwhite populations.[25] The point estimates for both white and non-white births are similar to the estimates of effects on average birth weight for any race, and results by race are not statistically distinguishable.

Next, we look for heterogeneity by the timing of grant receipt. If states wrote their priority lists to address the most severe pollution problems first, we would expect grants from the first few years of the CWA to have the largest effect on infant health. This is especially true if we think there is a convex relationship between pollution and health.

We address this in columns 3 and 4 of Table 2.5. In column 3, we drop all observations from facilities that received a grant after 1976 and re-estimate equation 2.3, and in column 4 we drop all observations from facilities that received a grant in or before 1976. The results are similar, so there is little evidence of heterogeneous effects by grant timing.

2.6.2 Mechanisms

2.6.2.1 Public Water

Public water systems, including those that draw from a surface water source such as a lake or river, often violate health-based water quality standards, and there is evidence that these violations impact infant and child health (Currie et al., 2013; Grossman and Slusky, 2019; Marcus, 2021). A report by the US Geological Survey (USGS) found that more than one in five source-water samples from public water systems contained one or more contaminants at concentrations dangerous to human health. In an analysis of matched water samples from 94 water sources and their associated public water systems, the same organic contaminants detected in source water consistently appeared at similar concentrations in drinking water after treatment (Toccalino and Hopple, 2010), suggesting that policies targeting improvements in surface water quality may have important impacts on health. With over 70 percent of community water system users receiving drinking water from a surface water source as of 1970 (Dieter, 2018), improvements in surface water pollution through CWA grants may have reduced exposure through improving public drinking water quality.

If reductions in contaminated public drinking water are driving health improvements, we would expect to find larger effects in areas that source public water from surface water rather than groundwater, as CWA

[25] The sample is restricted to counties where both the white and non-white average birth weight is calculated from 5 or more births. This ensures that we are making comparisons that rely on the same set of counties, in which there are sufficient individuals in both racial groups, rather than making comparisons between majority white and majority non-white communities. Results are not sensitive to this sample restriction.

grants directly affected surface water quality. We use USGS water use data from Solley et al. (1988) to divide our sample into counties that had any public water system that drew from surface water in 1985, and counties whose public water systems drew exclusively from ground water.[26]

We show that our results are driven by counties that had some public water systems that drew from surface water sources in Table 2.6. Column 1 of Table 2.6 re-estimates equation 2.4 on facilities whose downstream counties had some public water systems that drew from surface water sources, while column 2 estimates the same specification on facilities whose downstream counties' public water systems drew from groundwater exclusively. CWA grants significantly increased birth weight for counties where some drinking water is sourced from surface water, but there is no significant effect among counties that provide drinking water exclusively from groundwater sources. In fact, the point estimate is negative for these counties.[27]

We disaggregate these results further in Table 2.7 by estimating a triple difference where the first difference comes from where and when CWA grants were distributed, the second difference comes from if a birth occurred up or downstream from a wastewater treatment facility, and the third difference comes from whether downstream public water systems drew from surface or groundwater. Panels A and B estimate this triple difference on a sample of non-compliant facilities. We see strongly significant increases in birth weight and marginally significant decreases in the probability of low birth weight in areas that drew from surface water sources. Our estimates for areas that drew exclusively from groundwater are statistically insignificant and wrong-signed, and the birth weight effect in areas that drew from surface water is statistically greater than the effect in areas that only drew from groundwater. In Panels C and D, we re-estimate these specifications on samples of compliant facilities. These estimates can be thought of as a placebo test since these facilities experienced no improvement in downstream water quality. We find no significant effects of treatment in areas whose community water systems drew from either surface or ground water sources, as we would have expected. This suggests that our results are almost completely driven by counties that are downstream from non-compliant facilities in which some public water systems draw from surface water.

We provide further evidence that the effect of CWA grants on birth weight is driven by reduced contamination of publicly provided water in Table 2.8. Rather than defining the treated population as the percent of a county's population living within 1 mile of a treated waterway, we instead leverage information on the location of community water system service areas to define the treated population as the percent of the county's population served by a public drinking water system that is near a treated waterway. We calculate this using maps of public water supply areas from 8 states (see Section 2.6.5.3 for details on this data). Due to reduced

[26] We use data from 1985 because it is the earliest year for which information on county level water usage is available. While water service areas and county borders do not always perfectly align, community water systems generally serve areas no larger than counties (USEPA, 1997).

[27] Columns 6 and 7 of Table 2.4 suggest that communities served by surface and groundwater systems serve similar populations.

sample size, our results from this specification are less precise than our main results, however, the effects on both birth weight and probability of low birth weight are right-signed, and the effect on birth weight is marginally significant.

In this table, we re-estimate equation 2.2 defining pct_{cy} as the percent of the population that is served by a public drinking water system that is near a treated waterway. We calculate this using maps of public water supply areas from 8 states (see Section 2.6.5.3 for details on this data). Due to reduced sample size, our results from this specification are less precise than our main results, however, the effects on both birth weight and probability of low birth weight are right-signed, and the effect on birth weight is marginally significant.[28]

2.6.2.2 Recreation

While improved drinking water quality appears to be one channel through which grants improved infant health, improved surface water quality could also affect maternal health through water recreation. This channel could lead to improvements in health directly by reducing contact with contaminated water, or indirectly by making mothers more likely to exercise with a swim or a walk along a waterway. If recreational exposure is a primary channel through which these health effects occur, we might expect to find larger health improvements in states with more water-related recreation. While we do not observe water-related recreation activities directly, we can proxy for these activities using state-level per capita water recreation spending from the US Bureau of Economic Analysis.[29]

First, we test this channel in Table 2.10 by separately estimating equation 2.2 on sub-samples defined by terciles of state-level per capita water recreation spending. While we find the largest and most significant results in states in tercile 3, which had the highest water recreation spending, the confidence intervals for all three terciles overlap and estimates for each tercile are not statistically distinguishable in the pooled sample for average birth weight.

Next, in Table 2.11, we estimate our triple difference specification for each tercile of water recreation spending. Because the spending data is at the state level, we drop observations from the 889 facilities that have up and downstream counties in different states. In this specification, the middle tercile of states is driving our results, however, we still cannot statistically distinguish the point estimates across the three terciles.

Due to data limitations in our measurement of water recreation, we are not able to draw strong conclusions as to whether recreational activities contribute to our main findings. This is perhaps not surprising as recent research has highlighted the difficulty in accurately capturing water's recreational benefits (Kuwayama et al.,

[28] The estimates in Table 2.8 are smaller than those from the full sample. We re-estimate equation 2.2 on the sample of states that we have public water supply data for in Table 2.9, which shows that we obtain similarly smaller results on this reduced sample with our main specification.

[29] We focus on total spending for "Boating/Fishing" from 2012 to 2016, which includes canoeing/kayaking, fishing, sailing, and other boating. While data from the 1970's is not available, it is unlikely that cross-sectional variation in per capita recreational spending is changing much over time.

2018).

2.6.3 Mortality

Using data from National Center for Health Statistics (1988b), we re-estimate equation 2.4 with mortality as the dependent variable in Table 2.12. Columns 1-6 presents estimates from different age bins, and column 7 estimates the effect on mortality of child bearing age women. While these estimates are noisy, we find no significant effect of treatment on mortality for any group.

2.6.4 Robustness

2.6.4.1 Robustness to Distance Downstream

In the main text, we follow Keiser and Shapiro (2019a) and the EPA (USEPA, 2001) by defining a waterway as treated if it is 25 miles downstream from a wastewater treatment facility. We show that our results are not sensitive to this choice by re-estimating equation 2.4 defining treated waterways as those either 5 or 10 miles downstream from a treated facility in Table 2.13. The results are similar to those presented in Section 2.4.

2.6.4.2 Robustness to Buffer Selection

We expect the health effects of improved surface water quality to be concentrated near treated waterways. The exposed fraction of a county's population depends on the number of individuals living near a treated waterway, so we use census block population density data from the US Census Bureau to scale our results by the percent of a county's population living near a treated waterway. In our main results, we scale by the percent of a county's population living within a mile of a treated waterway, but Table 2.14 shows that our results are robust to scaling by the percent of a county's population living within other bandwidths around treated waterways by estimating equation 2.4 with different bandwidths. In column 1, this bandwidth is .5 miles, and in column 2, it is 1.5 miles.

2.6.4.3 Stacked Difference-in-Difference

Since we estimate two way fixed effects regressions, our results in the main text are an average of comparisons of (1) newly treated facilities relative to never-treated facilities, (2) newly treated facilities relative to facilities that have not yet been treated, and (3) newly treated facilities relative to already-treated facilities. When treatment effects are dynamic, the third type of comparison can be wrong signed (Goodman-Bacon, 2021a). We can get estimates that do not include comparisons of newly treated facilities relative to already-treated facilities, and explore if our results are driven by comparisons of treated units to not-yet-treated units or never-treated units by re-organizing our data into "stacks".

A stack is defined by a treatment cohort, that is, a group of facilities that received their first grants in a given year (e.g. every facility that received its first grant in 1974). Each stack contains observations from every facility in a treatment cohort, which are labeled as treated in that stack, and a set of controls that consist of either units that were treated at least eight years in the future, or all never-treated facilities. We can then estimate the following stacked difference-in-difference:

$$Y_{py} = \gamma^{stacked} pct_{py} + \alpha_{ps} + \alpha_{sy} + \varepsilon_{psy} \tag{2.5}$$

p indexes facilities, y indexes years, and s indexes stacks. Facility-by-stack fixed effects, α_{ps}, are analogous to a unit fixed effect in our regressions in the main text. Year-by-stack fixed effects, α_{sy}, ensure that we are only making comparisons within stacks, so our coefficient will not be identified off of comparisons of newly treated facilities relative to already-treated facilities.

We present estimates of equation 2.5 in Table 2.15. In column 1, the control group is not-yet-treated facilities. In column 2, it is never-treated facilities. In column 3, both never treated and not-yet-treated facilities are in the control group. We find significant effects on birth weight and the probability of low birth weight regardless of which control group we use. The effects are much larger when we compare treated units to never treated units, but since there are fewer never treated facilities than treated facilities, and since our two way fixed effect estimator averages these two effects together, our main results are closer to the results in column 1 than those in column 2.

2.6.4.4 Binary Treatment

Our main results define treatment with a continuous measure, so our results are identified in part off of comparisons between counties where a large proportion of the population is treated relative to counties where a small proportion is treated. Since we expect birth outcomes to improve homogeneously as more of the population becomes treated, there is nothing wrong with using this variation (Goodman-Bacon, 2021a), however, we can also define treatment in a binary way with a dummy variable that turns on after a county is downstream from a treated facility.

We first estimate the following event study

$$Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} + \beta X_{cy} + \alpha_c + \alpha_y + \varepsilon_{cy} \tag{2.6}$$

We present estimates of equation 2.6 with average birth weight and the probability of low birth weight in Figures 2.10 and 2.11. The shapes of these event studies are similar to those in the main text.

When we define treatment with a dummy variable, we can deal with the problems caused by dynamic treatment effects discussed in Section 2.6.4.3 in a more sophisticated way. To summarize these event studies, we use Callaway and Sant'Anna (2020a) to estimate treatment effects in Table 2.16.

Defining treatment in a binary way at the county level includes many untreated births, so these estimates are somewhat smaller and less significant than those in the main text, however, they are of the same sign as our main results, and the birth weight estimate is still marginally significant despite this attenuation.

2.6.4.5 Flow Rate, Population Served, and Non-Treatment Technology Modifications

In our triple difference specification, we interact treatment with a variable that indicates whether plants were compliant with new treatment technology standards when the CWA came into effect. Compliance is strongly correlated with heterogeneity in the effect of grants, but there could be other attributes correlated with grant effectiveness. To argue that the difference in grant effectiveness is due to differences in compliance, we interact treatment with measures of these other characteristics in Table 2.17 by estimating equation 2.7.

$$\Delta Y_{py} = \gamma pct_{py} + \eta pct_{py} * t_p + \pi pct_{py} * Interact_p + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py} \tag{2.7}$$

In column 1, the interaction term is the flow rate of the receiving facility measured in millions of gallons per day. In column 2, it is the total population served by the facility. In column 3, it is a dummy variable that equals one for facilities that indicated that they would use grant money to pay for non-treatment technology related upgrades in the 1972 CWNS. Column 4 includes all of these interactions in one equation.[30] All other variables are defined analogously to those in equation 2.3.

The coefficients on all three of the interaction terms are insignificant, and all three are wrong signed in columns 1 through 3, showing that facility size, the size of the population served, and non-treatment technology upgrades are not driving the heterogeneity in our estimates. This is further evidence that improvements in downstream infant health are driven by upgrades to treatment technology.

2.6.4.6 Unbalanced Event Study

In the main text, we look at effects up to eight years after treatment. Since we bin observations from greater than 8 years after treatment, we are only estimate balanced event study coefficients. We look at a longer post period by re-estimating the results in Figure 2.2 without binning these unbalanced endpoints in Figure 2.12. Since only early treated counties contribute to later event study coefficients, they should be interpreted with caution, however, these results suggest that the effect of CWA grants on infant health flattened out by 10 years

[30] We do not have data on these interaction terms for all facilities.

after treatment, consistent with projects taking up to 10 years from grant application to project completion (USEPA, 2002).

2.6.5 Additional Data Details

2.6.5.1 County Changes

Births records in NCHS data contain information on birth location at the county level. Several counties split or combined during our study period. Following Forstall (1995), we re-combine all counties that split or merged between 1968 and 1988. Changes are noted in Table 2.18.

2.6.5.2 Changes in Reported Sample

Data in years prior to 1972 constitutes a 50 percent sample of all births in the US. Years after 1972 contain information on every birth in the US from some states, and a 50 percent sample from the remaining states. Six states had full sample data in 1972, and all States and the District of Columbia had full sample data by 1985. Table 2.19 details the first year in which each state reported full sample data.

Our main results are weighted by total number of births in a county. Total births for observations from state-years reporting a 50 percent sample of births are defined as the number of observations from that county-year multiplied by two.

Changes from half to full sample often occurred around the same time as treatment. To be certain that our results are not driven by this change, we take a 50 percent sample of births from state-years that reported full sample data and re-estimate the results in Figure 2.2 on this sample in Figure 2.13. We then re-estimate the results presented in Table 2.3 on this sample and report the results in Table 2.20, which are similar to those reported in Section 2.4.

2.6.5.3 Public Water Supply Data

Data from each state comes from different years and reflects different water sources. Data from each state is described below.

Arkansas

Arkansas data is from the Arkansas GIS office, and is a comprehensive geographic database of water utilities and services in the Arkansas public water system. A visual aid of water system boundaries overlaid on current digital aerial photography, associated road names, and landmarks, were verified by representatives of ADH to confirm the accuracy of the boundaries. First published in 2013, these maps were last updated in 2019 (Arkansas GIS Office, 2013).

Arizona

Arizona data is maintained by the Arizona Department of Water Resources (ADWR) and reflects community water systems as of 2020. To determine the service area, ADWR utilized primary data provided directly from the water system (i.e. PDF, shapefile, verbal definition). If primary data was unavailable, secondary data (i.e. Certificate of Convenience and Necessity (CCN), Census Designated Place shapefile from U.S Census Bureau) was utilized to determine service area boundaries (Arizona Department of Water Resources, 2020).

Connecticut

Cennecticut public water supply maps are maintained by the Connecticut State Department of Health (Connecticut State Department of Public Health, 2020).

Kansas

Kansas public water maps are maintained by the The Kansas Water Office (KWO) and reflect public water supplies as of 2007 (Kansas Water Office, 2020).

New Jersey

New Jersey data comes from the Division of Science, Research, and Technology (DSRT) at the New Jersey Department of Environmental Protection (NJDEP). The maps shows all systems that piped water for human consumption to at least 15 service connections used year-round, or regularly served at least 25 year-round residents in 1998 (New Jersey Department of Environmental Protection, 2004).

North Carolina

North Carolina data comes from the NC Dept. of Environmental Quality, Division of Water Resources, Public Water Supply Section (PWSS), and contains maps of public water supply from 2017 (North Carolina Department of Environmental Quality, 2017).

Pennsylvania

Pennsylvania maps show all areas served by a community water supply system that serves at least 15 service connections or 25 year-round residents, such as manufactured housing communities, municipal water systems, personal care homes and housing developments.

The locations were digitized from maps submitted with Annual Water Supply Report for 2000, 2001, 2002 and 2003 (Pennsylvania Spatial Data Access, 2015).

Texas

Texas maps, maintained by the Texas Commission on Environmental Quality, show approximate relative locations of public water supply areas current to 2020 (Texas Commission on Environmental Quality, 2020).

2.6.5.4 Data on Wastewater Treatment Facilities

We begin with grant data from the EPA's Grant Information Control System, which we obtained through a Freedom of Information Act request. This data contains information on the year that the EPA distributed each grant, which municipality received the grant, the specific wastewater treatment facility the grant was designated for and the amount distributed. Keiser and Shapiro (2019a) uses the same data, and Appendix Section B.4 of Keiser and Shapiro (2019a) demonstrates its accuracy.

The 33,429 grants in our sample exclude grant records that do not include a specific facility code, as it is unclear to what extent these grants were precisely for wastewater treatment plants. We also drop grant records that are missing information on when they were distributed, which further restricts our sample to 29,898 grants.

We define whether a facility was in compliance with the CWA's capital mandate using the 1972 Clean Watershed Needs Survey, which we merge to our grant data with a unique facility code. The CWNS is an assessment of the capital investment that publicly-owned wastewater treatment facilities required to come into compliance with the Clean Water Act, and contains information on which community the facility serves, the number of residents served, the total wastewater flowing through the facility, the treatment technology currently in place, whether the facility needs to meet standards higher than the EPA's secondary treatment requirement, and whether they are currently in compliance with these requirements. This data was provided to us by the EPA's CWNS team, and is the same data that Jerch (2018) uses to define compliance with the CWA's capital mandate.

We use a facility's answer to Question 21 on the CWNS questionnaire to define compliance. Question 21b asks if a facility needs to meet treatment technology requirements that are more stringent than the EPA's secondary treatment requirement. Question 21c then asks whether a facility is currently in compliance with both the EPAs secondary treatment mandate and any higher mandates. We define facilities that answered "yes" on question 21c as "compliant", and those that answer no as "non-compliant". This defines facilities that satisfied the CWA's capital mandate when the CWA came into effect but did not satisfy more stringent state standards as non-compliant. When we use counties up and downstream from compliant facilities as an additional control group, we want to capture the effect of grants that were not bound by any capital mandate, so we do not want to define facilities that were still required to make upgrades as compliant, even if they are using secondary treatment.

Note that many facilities installed tertiary treatment after the CWA came into effect (USEPA, 2000). This increase was likely driven by municipalities bound by state standards or compelled by lawsuits to make upgrades beyond secondary treatment.

2.6.6 Additional Figures and Tables

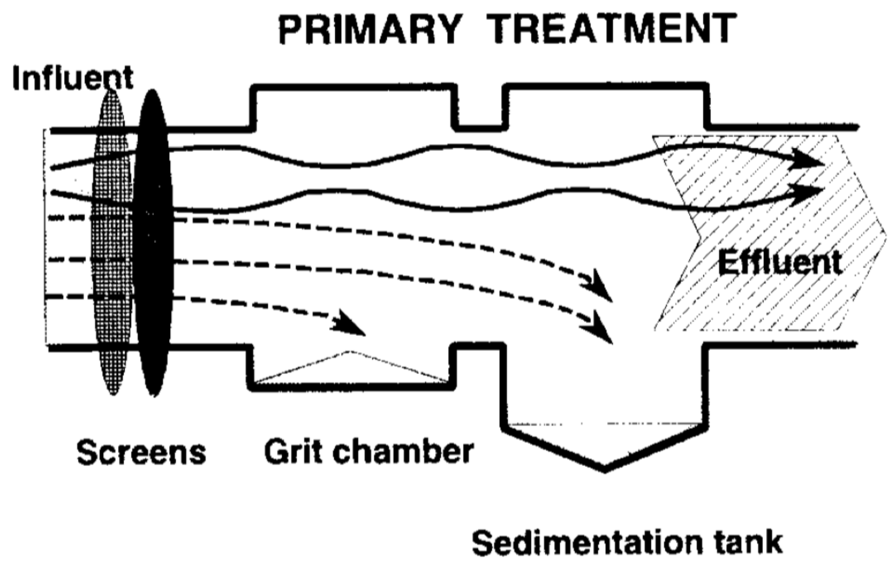


Figure 2.4: Primary Treatment Technology

Source: USEPA (1998)

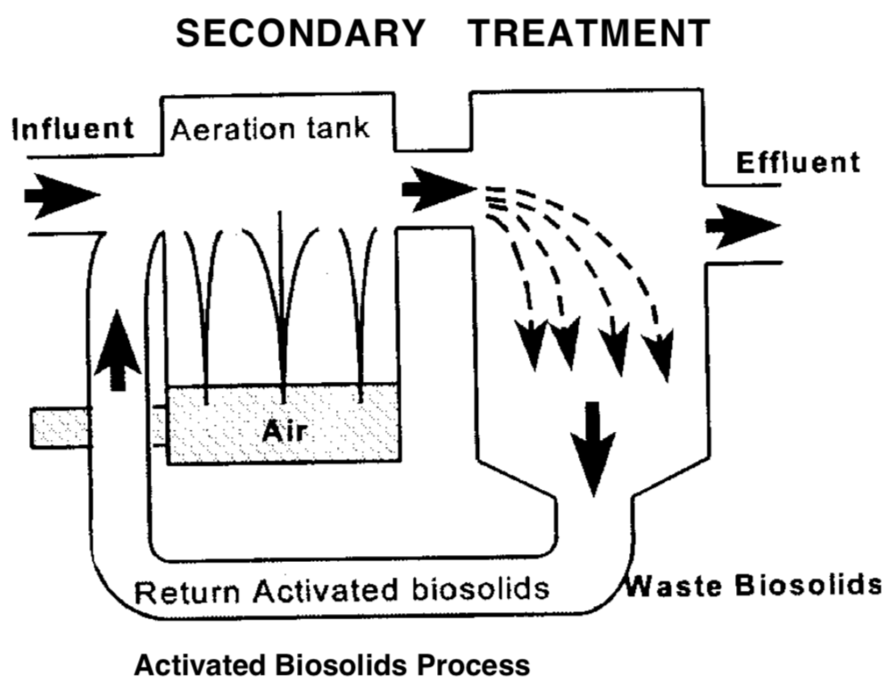


Figure 2.5: Secondary Treatment Technology

Source: USEPA (1998)

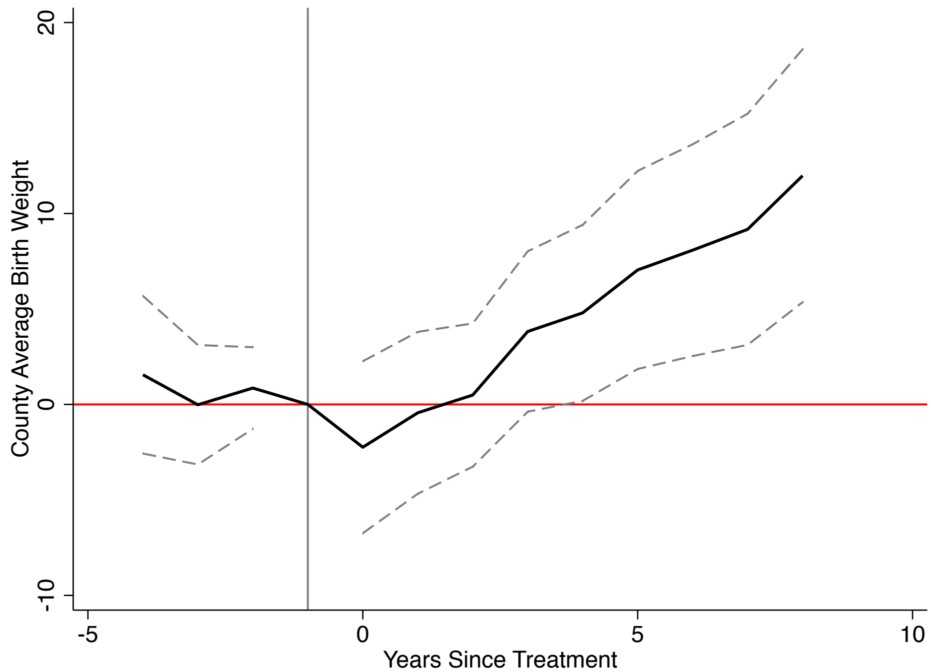
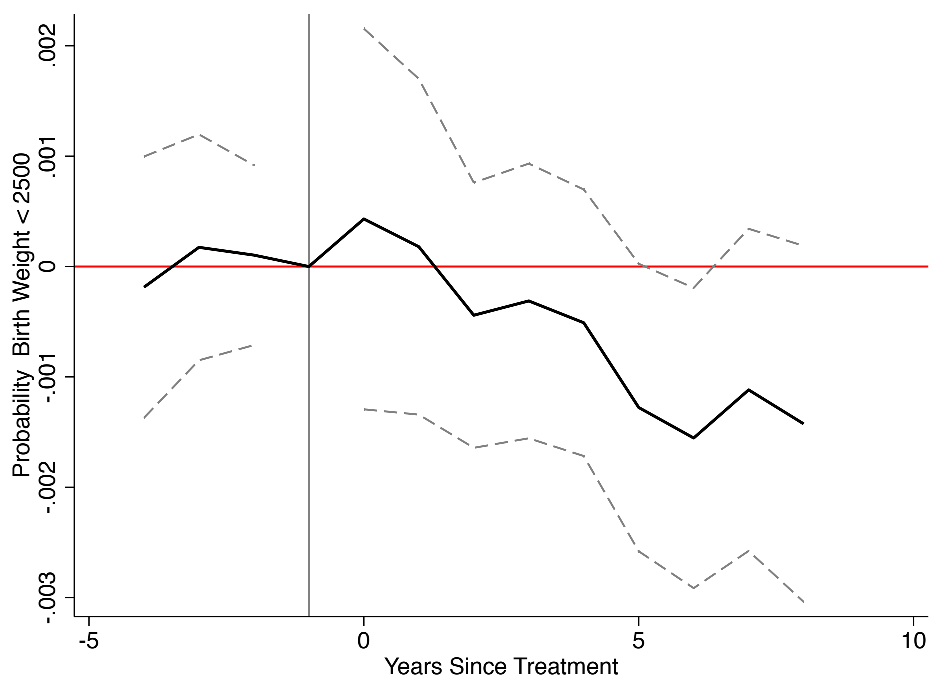


Figure 2.6: Birth Weight Downstream from Grant Facilities

Notes: This figures plot the π_t and γ_t from estimating $Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} * pct_{cy} + \beta X_{cy} + \alpha_c + \alpha_y + \varepsilon_{cy}$. pct_{cy} is a continuous variable that takes values from zero to one, and indicates the percent of county c's population living within a mile of a treated waterway in year y. The model includes county and year fixed effects, α_c and α_y respectively, as well as controls for the percent of a county's births of a given birth order, and county averages of mother's age and race and child gender. The estimates are weighted by total number of births in a county-year. The dependent variable is the the average birth weight in county c in year y.



=

Figure 2.7: Probability of Low Birth Weight Downstream from Grant Facilities

Notes: This figure re-creates Figure 2.6 with the probability of low birth weight as the dependant variable.

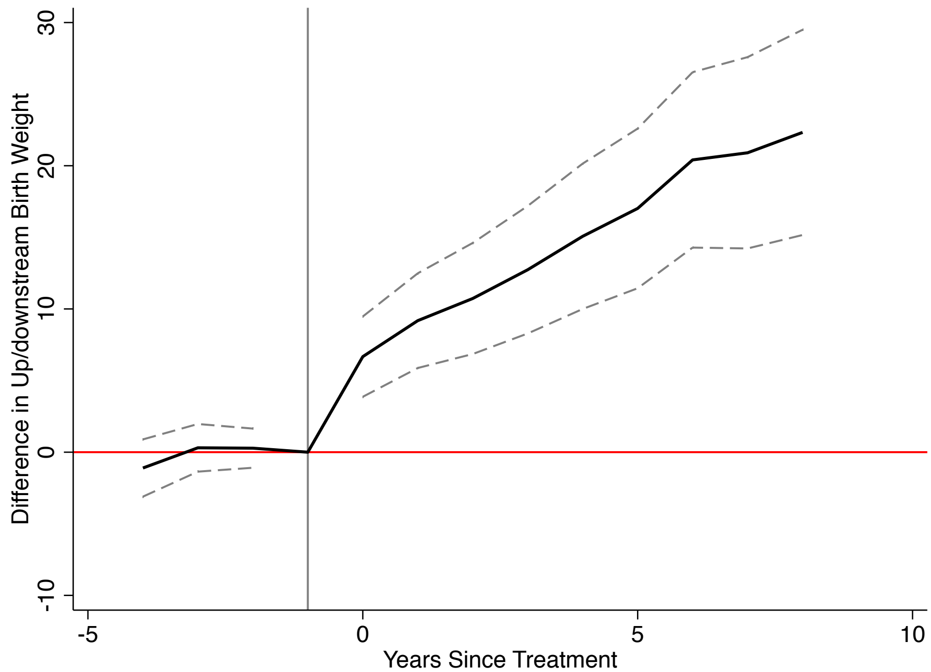


Figure 2.8: Difference in Birth Weight Up and Downstream from Grant Facilities

Notes: These figure plot the π_t and γ_t from estimating $\Delta Y_{py} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_p^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_p^* = t\} * pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py}$. The dependent variable is the difference in birth weight between up and downstream counties in year y.

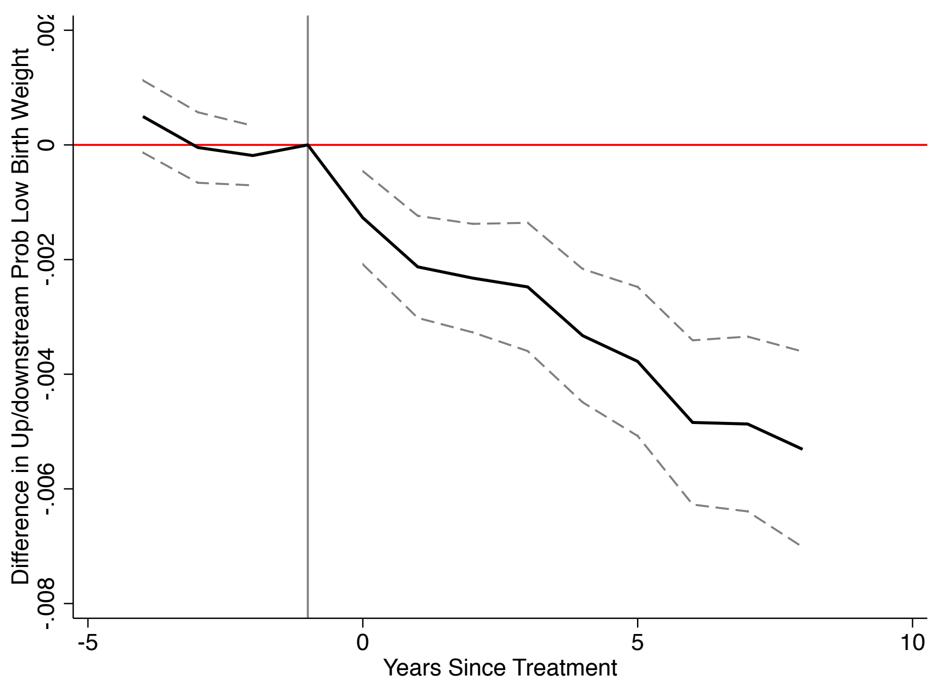


Figure 2.9: Difference in Probability of Low Birth Weight Up and Downstream from Grant Facilities

Notes: This figure re-creates Figure 2.8 with the difference in the probability of low birth weight between up and downstream counties in year y as the dependant variable.

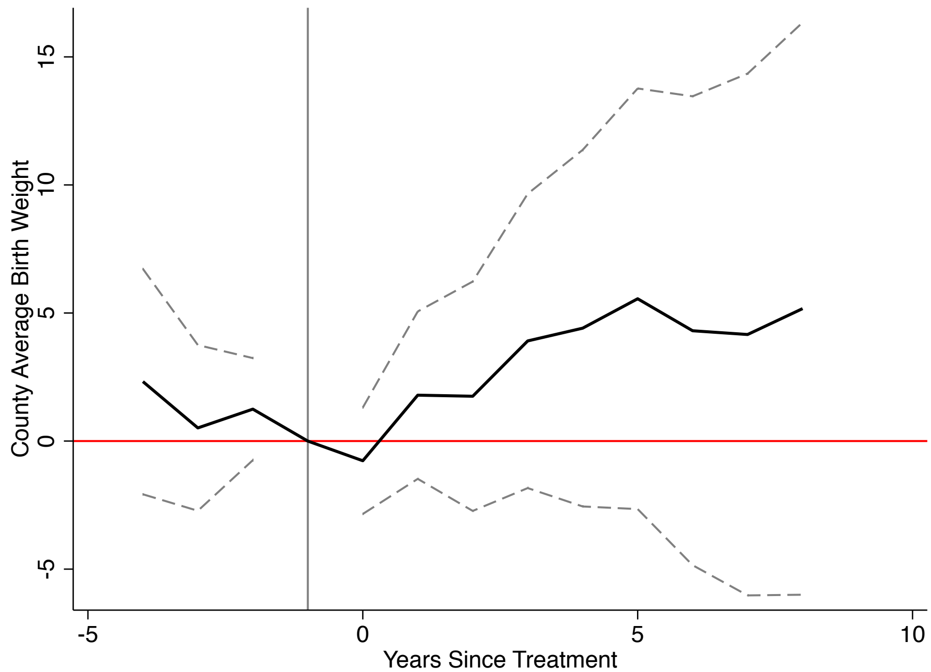


Figure 2.10: Birth Weight Downstream from Grant Facilities (Binary Treatment)

Notes: These figures plot the π_t and γ_t from estimating $Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} + \beta X_{cy} + \alpha_c + \alpha_y + \varepsilon_{cy}$. Regressions are weighted by the total number of births in county c in year y. The dependent variable is the the average birth weight in county c in year y.

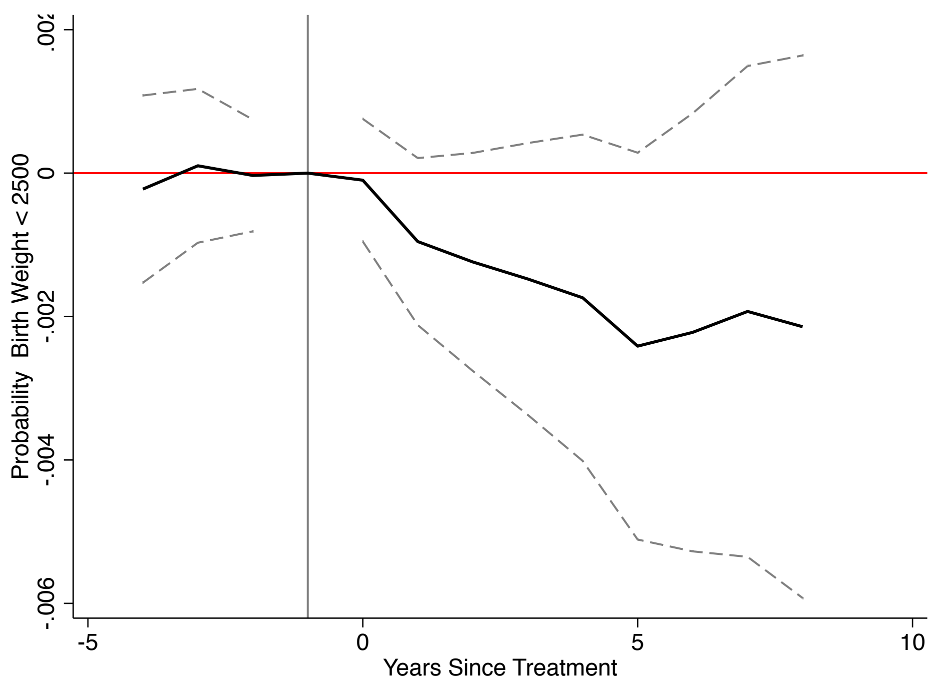


Figure 2.11: Probability of Low Birth Weight Downstream from Grant Facilities (Binary Treatment)

Notes: This figure re-creates Figure 2.8 with the difference in the probability of low birth weight as the dependant variable.

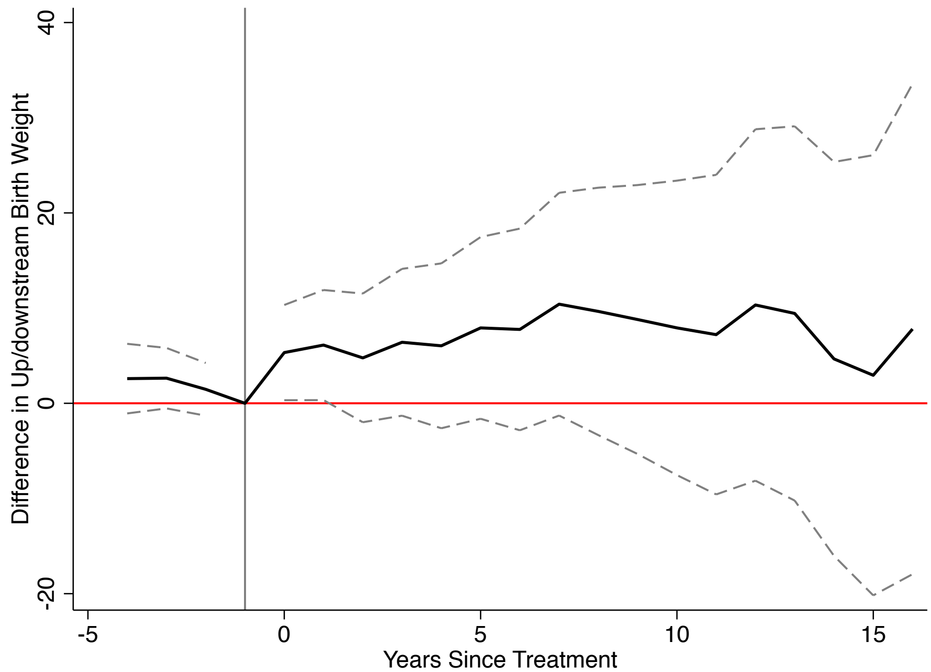


Figure 2.12: Birth Weight Triple Difference (Long Post)

Notes: These figures plot the θ_t and η_t from estimating $\Delta Y_{py} = \sum_{t=-5}^{-2} \theta_t 1\{y - y_p^* = t\} * t_p + \sum_{t=0}^{16} \eta_t 1\{y - y_p^* = t\} * pct_{py} * t_p + \sum_{t=-4}^{-2} \pi_t 1\{y - y_p^* = t\} + \sum_{t=0}^{16} \gamma_t 1\{y - y_p^* = t\} * pct_{py} + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py}$. All variables are defined analogously to those in Figure 2.1. The dependent variable is the difference in birth weight between up and downstream counties in year y.

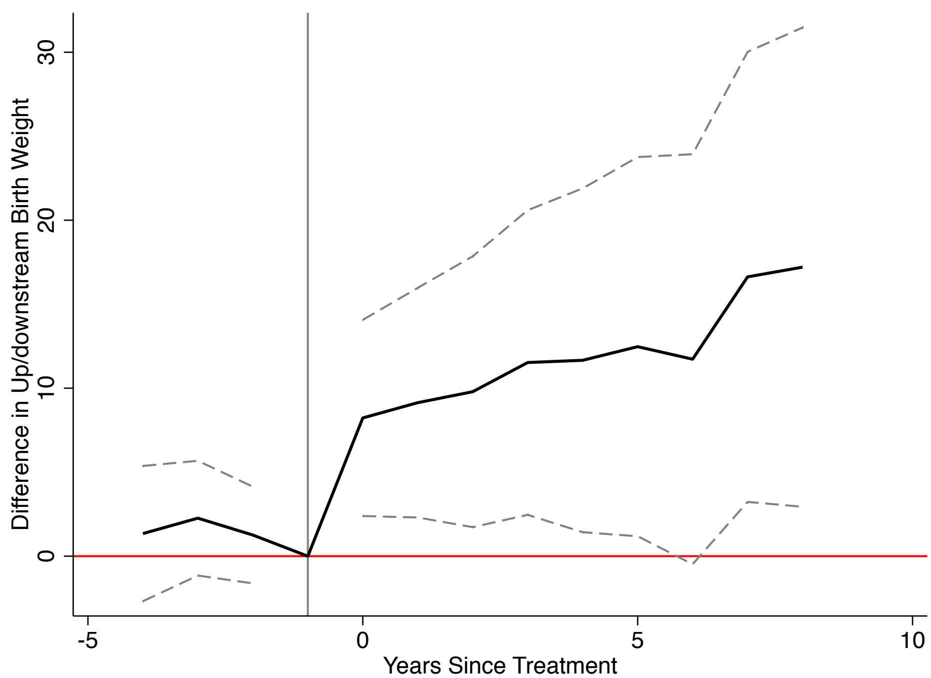


Figure 2.13: Birth Weight Triple Difference (Random Sample)

Notes: This Figure re-estimates the results in Figure 2.2 after taking a fifty percent random sample of births that occurred in state-years that reported a full sample of births. The years that each state switched from a 50 percent sample to a full sample of births are detailed in Table 2.19.

Table 2.4: Summary Statistics

	(1) Full Sample	(2) Downstream	(3) Upstream	(4) Non-Compliant	(5) Compliant	(6) Surface	(7) Ground
Birth Weight	3279.61	3277.83	3297.25	3279.70	3279.37	3275.67	3296.68
Probability Birth Weight < 2500	.078	.079	.074	.078	.077	.078	.077
Non-White	.166	.170	.115	.155	.193	.161	.185
Age of Mother	24.58	24.58	24.62	24.66	24.39	24.63	24.40
Education of Mother	11.83	11.83	11.83	11.87	11.65	11.86	11.72
Birth Order	2.40	2.39	2.42	2.42	2.34	2.37	2.52
Observations	1788138	1571197	206017	1300614	487524	1452552	335586

Notes: This table presents the mean of birth weight, the probability of low birth weight, the percent of non-white births, average age and education of mothers, and average birth order for all counties, births in counties that were ever downstream from a facility that received a CWA grant, counties that were ever upstream from a facility that received a CWA grant, counties up or downstream from non-compliant facilities, counties up or downstream from compliant facilities, counties that had at least some public water systems that drew from surface water, and counties that used exclusively ground water. These means are calculated using individial birth data from 1970, two years before the CWA came into effect.
Source: National Center for Health Statistics (1988a)

Table 2.5: Heterogeneous Effects

	(1) White	(2) Non-White	(3) Early Grants	(4) Later Grants
Pct Pop 1 Mile X Non-Compliant	11.37***	14.32	14.04**	11.95**
	(3.872)	(10.89)	(6.526)	(5.366)
Demographic Controls	X	X	X	X
Unit and Year Fixed Effects	X	X	X	X
Collapsed to Facility Level	X	X	X	X
Observations	35406	35406	51639	31080

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the equation 2.4 on sub-samples of the population. Columns 1 and 2 divide the sample by race and only include counties that had a sizeable nonwhite population, and columns 3 and 4 divide the sample by grant timing.
Source: National Center for Health Statistics (1988a)

Table 2.6: Effects by Public Water Source

	Surface Water (1)	Ground Water (2)
Panel A	County Average Birth Weight	
Pct Pop 1 Mile X Non-Compliant	8.893**	-5.137
	(3.579)	(8.252)
Panel B	Probability Birth Weight < 2500 grams	
Pct Pop 1 Mile X Non-Compliant	-0.000952	0.000132
	(0.000845)	(0.00198)
Demographic Controls	X	X
Unit and Year Fixed Effects	X	X
Collapsed to Facility Level	X	X
Observations	67032	15288

95% confidence intervals in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the specification in column 7 of Table 2.3 on sub-samples of counties that had some public water systems that draw from surface water and counties whose public water systems only draw from groundwater.
Source: National Center for Health Statistics (1988a); Solley et al. (1988)

Table 2.7: Public Water Source Triple Difference

	Surface (1)	Ground (2)	DDD (3)
Panel A. Non-compliant	County Average Birth Weight		
Pct Pop 1 Mile	10.15***	-7.879	-7.879
	(2.154)	(6.327)	(6.115)
Pct Pop 1 Mile X Surface			18.03***
			(6.451)
Observations	30009	4200	34209
Panel B. Non-compliant	Probability Birth Weight $<$ 2500 grams		
Pct Pop 1 Mile	-0.000872*	0.00103	0.00103
	(0.000485)	(0.00150)	(0.00145)
Pct Pop 1 Mile X Surface			-0.00190
			(0.00152)
Observations	30009	4200	34209
Panel C. Compliant	County Average Birth Weight		
Pct Pop 1 Mile	3.111	3.110	3.110
	(2.025)	(3.836)	(3.774)
Pct Pop 1 Mile X Surface			0.000404
			(4.260)
Observations	37023	11088	48111
Panel D. Compliant	Probability Birth Weight $<$ 2500 grams		
Pct Pop 1 Mile	-0.000333	-0.00183	-0.00183
	(0.000534)	(0.00120)	(0.00115)
Pct Pop 1 Mile X Surface			0.00150
			(0.00126)
Observations	37023	11088	48111
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to Facility Level	X	X	X

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table describes the effects of Clean Water Act grants on birth outcomes depending on public water source. Column 1 estimates $\Delta Y_{py} = \gamma pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py}$ for facilities whose downstream counties had some public water systems that drew from surface water, and column 2 re-estimates this specification for counties whose public water systems only drew from groundwater. Column 3 estimates the associated triple difference: $\Delta Y_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * s_p + \beta X_{py} + \phi X_{py} * s_p + \alpha_y * s_p p + \alpha_p + \alpha_y + \varepsilon_{py}$ where s_p is a dummy variable that equals one for facilities with downstream counties that drew at least some drinking water from surface water sources. All regressions include demographic controls and unit and year fixed effects. Panels A and B run this analysis for non-compliant facilities, and Panels C and D repeat this analysis for compliant facilities as a robustness check. Average birth weight is the dependent variable in Panels A and C, and probability of low birth weight is the dependent variable in Panels B and D.
Source: National Center for Health Statistics (1988a); Solley et al. (1988)

Table 2.8: Using Public Water Supply to Define Treatment

	(1) Birth Weight	(2) Probability Birth Weight $<$ 2500
Pct Pop Public Water	4.705*	-0.000224
	(2.602)	(0.000954)
Demographic Controls	X	X
Unit and Year Fixed Effects	X	X
Collapsed to County Level	X	X
Observations	8463	8463

Standard errors in parenthesis, clustered at county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: In this table, we re-estimate the results in column 2 of Table 2.3 defining pct_{cy} as the percent of the population that is served by a public drinking water system that is near a treated waterway.
Source: National Center for Health Statistics (1988a)

Table 2.9: Main Results for States with Public Water Supply Data

	(1) birth weight	(2) prob bw $<$ 2500
pct pop 1 mile	2.242	-0.000626
	(3.195)	(0.00104)
demographic controls	X	X
unit and year fixed effects	X	X
collapsed to county level	X	X
Observations	8463	8463

Standard errors in pare thesis, clustered at county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: In this table, we re-estimate the results in column 2 of Table 2.3 on the eight states that we have public water supply data for.
Source: National Center for Health Statistics (1988a)

Table 2.10: Results by Per Capita Recreational Spending

	(1) Tercile 1	(2) Tercile 2	(3) Tercile 3
Panel A	County Average Birth Weight		
Pct Pop 1 Mile	0.163	4.701	15.19***
	(3.960)	(3.368)	(4.427)
Panel B	Probability Birth Weight < 2500 grams		
Pct Pop 1 Mile	-0.000429	-0.000453	-0.00220***
	(0.000766)	(0.000960)	(0.000837)
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to County Level	X	X	X
Observations	21147	20160	22617

Standard errors in parenthesis, clustered at county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the specification in column 2 of Table 2.3 on sub-samples defined by terciles of state water-related recreational spending. Counties in states with the lowest spending are in Tercile 1, while those in states with the highest spending are in Tercile 3.
Source: National Center for Health Statistics (1988a); Bureau of Economic Analysis (2017)

Table 2.11: Triple Difference by Per Capita Recreational Spending

	(1) Tercile 1	(2) Tercile 2	(3) Tercile 3
Panel A	County Average Birth Weight		
Pct Pop 1 Mile X non-compliant	1.231	14.76***	4.012
	(5.552)	(5.604)	(5.833)
Panel B	Probability Birth Weight < 2500 grams		
Pct Pop 1 Mile X non-compliant	0.00183	-0.00352**	0.000262
	(0.00123)	(0.00171)	(0.00141)
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to County Level	X	X	X
Observations	20748	19656	23247

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the specification in column 7 of Table 2.3 on sub-samples defined by terciles of state water-related recreational spending. Facilities in states with the lowest spending are in Tercile 1, while those in states with the highest spending are in Tercile 3.
Source: National Center for Health Statistics (1988a); Bureau of Economic Analysis (2017)

Table 2.12: Mortality Triple Difference

	(1) under 1	(2) 1-19	(3) 20-44	(4) 45-64	(5) 65-84	(6) 85+	(7) women 15-44
Pct Pop 1 Mile X Non-Compliant	0.389	10.11	-14.51	-3.723	-35.34	-19.66	1.607
	(10.22)	(10.26)	(24.77)	(20.17)	(43.11)	(24.78)	(5.157)
Demographic Controls	X	X	X	X	X	X	X
Unit and Year Fixed Effects	X	X	X	X	X	X	X
Collapsed to Facility Level	X	X	X	X	X	X	X
Observations	82320	82320	82320	82320	82320	82320	82320

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents (weighted) estimates from the following model: $\Delta Y_{py} = \gamma pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py}$. The dependent variable is the difference in mortality between counties up and downstream from facility p in year y. Columns 1-6 presents estimates from different age bins, and column 7 estimates the effect on mortality of child bearing age women.
Source: National Center for Health Statistics (1988b); Solley et al. (1988)

Table 2.13: Other Distances Downstream

	Non-Compliant (1)	Compliant (2)	DDD (3)
Panel A. 5 Miles Downstream	County Average Birth Weight		
Pct Pop 1 Mile	14.68***	6.358***	6.358***
	(2.800)	(2.125)	(2.125)
Pct Pop 1 Mile X Non-Compliant			8.326**
			(3.515)
Observations	35973	50379	86352
Panel B. 10 Miles Downstream	County Average Birth Weight		
Pct Pop 1 Mile	14.44***	6.167***	6.167***
	(2.783)	(2.113)	(2.113)
Pct Pop 1 Mile X Non-Compliant			8.278**
			(3.493)
Observations	35154	49413	84567
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to Facility Level	X	X	X

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents (weighted) estimates from the following model: $bw_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py}$. pct_{cy} is a continuous variable that takes values from zero to one, and indicates the proportion of downstream counties' populations that lived within a mile of a treated waterway in a given year. In Panel A, a waterway is considered treated if it is within 5 miles downstream from a facility that received a Clean Water Act grant. In Panel B, a waterway is considered treated if it is within 10 miles downstream from a facility that received a Clean Water Act grant.
Source: National Center for Health Statistics (1988a)

Table 2.14: Alternative Small Bandwidths

	(1) 25 Miles Downstream 0.5 Mile Buffer	(2) 25 Miles Downstream 1.5 Mile Buffer
	County Average Birth Weight	
Pct Pop 0.5 Miles	10.70**	
	(4.458)	
Pct Pop 1.5 Miles		6.621**
		(2.826)
Demographic Controls	X	X
Unit and Year Fixed Effects	X	X
Collapsed to Facility Level	X	X
Observations	82320	82320

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents (weighted) estimates from the following model: $\Delta Y_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \varepsilon_{py}$. pct_{py} is a continuous variable that takes values from zero to one, and indicates the proportion of downstream counties' populations that lived within some bandwidth of a treated waterway in a given year. In column 1, this bandwidth is .5 miles, and in column 2, it is 1.5 miles.

Table 2.15: Stacked Difference in Difference

	(1) Not Yet Treated	(2) Never Treated	(3) Both
Panel A	County Average Birth Weight		
Pct Pop 1 Mile	5.209**	26.96***	5.458**
	(2.531)	(3.997)	(2.524)
Panel B	Probability Birth Weight < 2500		
Pct Pop 1 Mile	-0.00134**	-0.00541***	-0.00139**
	(0.000554)	(0.000836)	(0.000552)
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to Facility Level	X	X	X
Observations	83580	63041	86088

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents results from estimating the following stacked difference in difference: $Y_{py} = \gamma^{stacked} pct_{psy} + \alpha_{ps} + \alpha_{sy} + \varepsilon_{psy}$. In column 1, the control group is facilities that will be treated at least 9 years in the future. In column 2, the control group is facilities that never receive a CWA grant. In column 3, both never treated and not-yet-treated units are in the control group. The dependent variable is the difference in birth weight between up and downstream counties in year y in Panel A, and the difference in the probability of being born weighing less than 2500 grams between up and downstream counties in year y in Panel B. Source: National Center for Health Statistics (1988a)

Table 2.16: Callaway and Sant'Anna (2020a) Estimates

	Birth Weight (1)	Probability Birth Weight < 2500 (2)
Grant X Downstream	4.85*	-0.0018
	(2.60)	(0.0032)
Observations	64239	64239

Standard errors in parenthesis, clustered at county level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents event study aggregations of group time average treatment effect estimates of the effect of being downstream from a facility that received a CWA grant on birth outcomes.
Source: National Center for Health Statistics (1988a)

Table 2.17: Other Interactions

	(1)	(2)	(3)	(4)
	County Average Birth Weight			
Pct Pop 1 Mile X Non-Compliant	6.464**	5.268**	5.389	6.736
	(2.957)	(2.614)	(3.843)	(4.493) [1em] Pct Pop 1 Mile
4.719*	7.304***	5.888	5.687	
	(2.664)	(2.315)	(3.918)	(4.402)
Pct Pop 1 Mile X Total Flow	-0.0263			0.0347
	(0.0198)			(0.0337)
Pct Pop 1 Mile X Population Served		-0.00000700		-0.0000165
		(0.00000582)		(0.0000108)
Pct Pop 1 Mile X Other Modification			-0.903	-2.871
			(6.745)	(7.081)
Demographic Controls	X	X	X	X
Unit and Year Fixed Effects	X	X	X	X
Collapsed to Facility Level	X	X	X	X
Observations	35049	45864	30597	24717

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $\Delta Y_{py} = \gamma pct_{py} + \eta pct_{py} * t_p + \pi pct_{py} * Interact_p + \beta X_{py} + \alpha_p + \alpha_y + \varepsilon_{py}$ In column 1, the interaction term is is the flow rate of the receiving facility measured in millions of gallons per day. In column 2, it is the total population served by the facility. In column 3, it is a dummy variable that equals one for facilities that indicated that they would use grant money to pay for non-treatment technology related upgrades in the 1972 CWNS. Column four includes all of these interaction terms. All other variables are defined analogously to those in equation 2.3.
Source: National Center for Health Statistics (1988a)

Table 2.18: County Code Changes

State fips	New County fips	Old County fips	Year	Note
4	12	27	1983	La Paz County, AZ split off from Yuma county
13	510	215	1971	The city of Columbus, GA became a consolidated city-county
29	186	193	N/A	Ste. Genevieve county, MO changed codes
32	510	25	1968	Ormsby County became Carson City
35	6	61	1981	Cibola County, NM split off from Valencia County
46	71	131	1979	Washabaugh County was annexed to Jackson County
51	83	780	1995	South Boston City rejoins Halifax County
51	510	13	N/A	Alexandria City/Arlington County
51	515	19	1968	Bedford City splits from Bedford County
51	520	191	N/A	Bristol City/Washington County
51	530	163	N/A	Buena Vista City/Rockbridge County
51	540	3	N/A	Charlottesville City/Albemarle County
51	560	75	N/A	Clifton Forge City/Alleghany County
51	590	143	N/A	Danville City/Pittsylvania County
51	630	177	N/A	Fredericksburg City/Spotsylvania County
51	660	165	N/A	Harrisonburg City/Rockingham County
51	670	149	N/A	Hopewell City/Prince George County
51	680	31	N/A	Lynchburg City/Campbell County
51	683	153	1975	Manassas City splits from Prince William County
51	685	153	1975	Manassas Park City splits from Prince William County
51	690	89	N/A	Martinsville City/Henry County
51	710		N/A	Norfolk City came from Norfolk County, which was ultimately combined into Chesapeake City
51	730	53	N/A	Petersburg City/Dinwiddie County
51	735	199	1975	Poquoson City splits from York County
51	740		N/A	Portsmouth City came from Norfolk County before it was Chesapeake City
51	750	121	N/A	Radford City/Montgomery County
51	770	161	N/A	Roanoke City/Roanoke County
51	775	161	1968	Salem City splits from Roanoke County
51	790	15	N/A	Staunton City//Augusta County
51	800	123	1974	Nansemond County merges into Suffolk City
51	840	69	N/A	Winchester City//Frederick County

Table 2.19: Sample Changes

State Name	State NCHS Code	State fips Code	First Full Sample Year
Alabama	1	1	1976
Arizona	3	4	1985
Arkansas	4	5	1980
California	5	6	1985
Colorado	6	8	1973
Connecticut	7	9	1979
Delaware	8	10	1985
Washington DC	9	11	1984
Florida	10	12	1972
Georgia	11	13	1985
Idaho	13	16	1977
Illinois	14	17	1974
Indiana	15	18	1978
Iowa	16	19	1974
Kansas	17	20	1974
Kentucky	18	21	1976
Louisiana	19	22	1975
Maine	20	23	1972
Maryland	21	24	1975
Massachusetts	22	25	1977
Michigan	23	26	1973
Minnesota	24	27	1976
Mississippi	25	28	1979
Missouri	26	29	1972
Montana	27	30	1974
Nebraska	28	31	1974
Nevada	29	32	1976
New Hampshire	30	33	1972
New Jersey	31	34	1979
New Mexico	32	35	1982
New York	33	36	1977
North Carolina	34	37	1975
North Dakota	35	38	1983
Ohio	36	39	1977
Oklahoma	37	40	1975
Oregon	38	41	1974
Pennsylvania	39	42	1979
Rhode Island	40	44	1972
South Carolina	41	45	1974
South Dakota	42	46	1980
Tennessee	43	47	1975
Texas	44	48	1976
Utah	45	49	1978
Vermont	46	50	1972
Virginia	47	51	1975
Washington	48	52	1978
West Virginia	49	53	1976
Wisconsin	50	55	1975
Wyoming	51	56	1979

Table 2.20: Triple Difference: Random Sample

	(1) Non-Compliant	(2) Compliant	(3) DDD
Pct Pop 1 Mile	12.38***	4.448**	4.448**
	(2.735)	(2.114)	(2.114)
Pct Pop 1 Mile X Non-Compliant			7.933**
			(3.456)
Demographic Controls	X	X	X
Unit and Year Fixed Effects	X	X	X
Collapsed to Facility Level	X	X	X
Observations	34188	48132	82320

Standard errors in parenthesis, clustered at facility level

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the specifications in Columns 5-7 in Panel A of Table 2.3 after taking a fifty percent random sample of births that occurred in state-years that reported a full sample of births.
Source: National Center for Health Statistics (1988a)

CHAPTER 3

Competitive Grants and Student Achievement: The Effect of Race to the Top Grants

Part of the American Recovery and Reinvestment Act of 2009 (ARRA), Race to the Top was a federal grant program in which the US Department of Education (ED) distributed $4.35 billion to 19 state governments. While Congress designed most of the education funding in the ARRA to stabilize education spending during the Great Recession, Congress designed Race to the Top grants to encourage long term policy changes. To incentivize states to enact new policies, ED distributed RTTT grants based on states' compliance with a number of education reforms.

This paper addresses two questions: what is the impact of RTTT grants on student achievement, and are these effects driven by extra funding, policy changes, or some combination of the two. I use a state level difference-in-differences design to show that RTTT policies cause increases in student achievement. These improvements are statistically larger in states that received grants, demonstrating that policy changes are more effective when paired with extra funding.

ED distributed RTTT grants in three rounds, but the program exhausted its funding in the first two rounds. The 12 states that received grants in the first two rounds (which I refer to as "early treated states") received grants over ten times larger than those distributed to the seven states that received grants in the third round of funding (which I refer to as "late treated states"). States that received RTTT grants in any round were more likely than non-RTTT states to implement RTTT policies (Boser, 2012; Howell, 2015), and policy adoption was similar across early and late treated states (Dragoset et al., 2016). The way in which ED distributed RTTT grants creates variation in the funding of these new policies across early and late treated states *without* creating variation in education policy.

Contemporary and retrospective reports suggest that early and late grant receipt was quasi-random. For example, a report by the Economic Policy Institute describes the grant allocation process as "subjective and arbitrary, more a matter of bias or chance than a result of ... states' superior compliance with reform policies" (Peterson and Rothstein, 2010). This supports making comparisons across early, late, and never treated states. Assuming that early and late treated states would have responded to policy changes in the same way in the absence of extra funding, comparing student achievement in early and late treated states before and after grant receipt will return the effect of additional funding on achievement. Comparing early treated states to never treated states will capture the effect of both extra funding and policy changes, and comparing late treated states to never treated states will capture the effect of policy changes.

Only one study, Dragoset et al. (2016), examines the effect of RTTT grants on state level student outcomes, and the results are inconclusive. There are several studies of the effects of individual policies promoted by RTTT. For example, Henry et al. (2015) uses a regression discontinuity design to show that a RTTT promoted policy in North Carolina in which the state department of education took over schools in the bottom five percent of achievement led to large improvements in test scores. Zimmer et al. (2017) finds mixed results when studying a similar program in Tennessee.

The results in Henry et al. (2015) and Zimmer et al. (2017) answer specific questions about the education production function, however, they do not inform grant design. While there is a large literature studying formula grants (Hines and Thaler, 1995; Inman, 2008), there is little research on the effectiveness of competitive grants outside of a few specific contexts, such as agricultural research grants to individuals (Huffman and Evenson, 2006) and grants to local governments to reduce wildfire risk (Cheng and Dale, 2020). Competitive grants have two major appeals: (1) many states that do not receive funding may still implement policies promoted by the program in an attempt to secure grant funding, and (2) the attached prestige of winning a grant might induce states to change policy in exchange for relatively small grants.

A better understanding of the advantages of competitive grants could allow donor governments to improve the benefit/cost ratio of many grant programs. A donor government will face the same costs to implement a grant program if they determine grant allocation based on policy implementation instead of a formula. For this reason, if competitive grants generate higher returns than formula grants, distributing grants competitively instead of distributing grants according to a formula may increase a grant program's benefit/cost ratio.

Measuring student achievement with the National Assessment of Educational Progress (NAEP), I show that proficiency rates improved in both early and late treated states, though these improvements were limited to states with low baseline achievement. This shows that RTTT policy changes led to improvements in achievement regardless of grant size, which suggests that competitive grant programs can cause the desired policy outcomes regardless of grant receipt. The improvements in student achievement are statistically larger in early treated states relative to late treated states, suggesting that RTTT policies were more effective when paired with additional funding.

3.1 Background

States had to apply to ED to receive RTTT grants. States' applications detailed their plans to change state education policies and documented any past progress in RTTT policy areas. ED scored these applications on a 500 point scale, and this score determined grant receipt. Figure 3.1 plots out each state's score along with the cutoff between early and late treated states. Scores change smoothly across the cutoff, and both early and late treated states earned substantially higher scores than most never treated states, suggesting that early and late treated states implemented similar policies.[1]

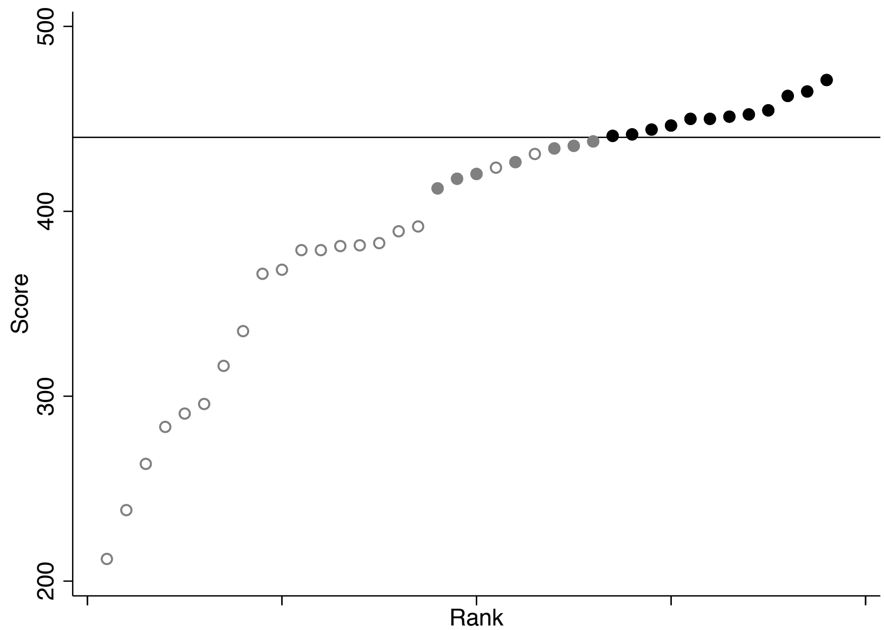


Figure 3.1: Early and Late Treated States' RTTT Applications Scored Similarly

Notes: This figure plots the score that each state received on their round 2 RTTT application (except for Tennessee and Delaware, which received grants in round 1, and for which I present the round 1 score) against their rank. Early treated states are represented by black dots, late treated states are represented by gray dots, and never treated states are represented by white dots. The line at 440 represents the cutoff score.

Reform policies fell into six broad categories: (1) improving state capacity; (2) implementing new student achievement standards; (3) building state data systems; (4) recruiting and retaining highly effective teachers; (5) turning around the lowest performing schools; and (6) removing restrictions on charter schools (Dragoset et al., 2016). States received a score on each of these categories. Figure 3.2 shows that, relative to early treated states, late treated states earned marginally lower scores in each of these categories.

Initially, ED planned on awarding large grants to all qualifying states. U.S. Education Secretary Arne Duncan said his department received "many more competitive applications than money to fund them" (Rundquist and Calefati, 2010), and that the department "funded as many states as [it] could [until it] ran out of money" (Paulson, 2011). Operating under the assumption that Congress would renew the program's funding, ED exhausted all of the funds made available to RTTT in grants to 12 states. ED petitioned Congress to provide an additional $1.35 billion for grants to the seven reaming qualifying states, but Congress approved less than $200 million dollars in new funding (Howell, 2015). Consequently, the 12 states treated in 2010 received

[1]Note that there are two never treated states with scores close to late treated states. These represent California and South Carolina, who withdrew from RTTT despite being eligible for small grants (Los Angeles Times, 2010). The fact that some states were eligible for small grants but did not accept them suggests that the states that accepted small grants took their education reforms seriously.

grants averaging about $328 million, or $57.1 per capita, while the seven runner up states received grants averaging $28.5 million, or $3.6 per capita.[2] Figure 3.3 maps out early and late treated states and Figure 3.4 shows average grant size in early and late treated states.

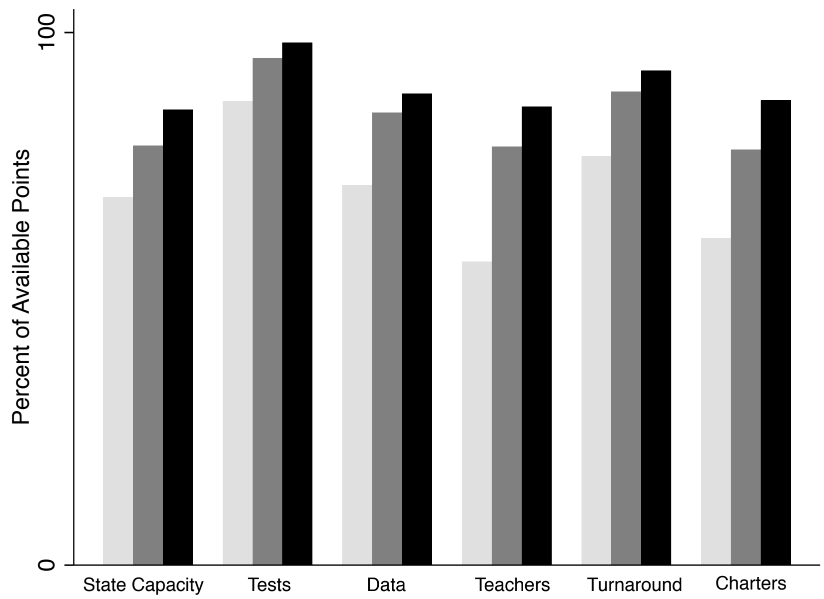


Figure 3.2: Category Specific Scores by Treatment

Notes: This figure dissagregates the scores from the Figure 3.1 into the six broad categories of RTTT reform policies. Average scores of never-treated states are presented in light gray, late treated states are in dark gray, and early treated states are in black.

Despite the difference in grant size associated with grant timing, there is evidence that early and late treatment was arbitrary. Dragoset et al. (2016), the official review of RTTT commissioned by ED, reported that "ED determined that [early and late treated states] alike had formed comprehensive plans to improve education policies and practices and had demonstrated the capacity for, and significant past progress in, implementing policies and practices in the program's topic areas".

As an example, New Jersey fell short of the score cutoff because of a single minor error in its application (Rundquist and Calefati, 2010). Since there was no appeals process, New Jersey received a grant of less than $40 million in 2011 as opposed to the $400 million it would have received if it had qualified for an early grant. A spokesperson for the New Jersey teachers union described this small grant as "less than a drop in the bucket... This is one-time money that will help set up some pilot programs... It eludes me how this will have a lasting impact" (Calefati, 2011).

[2]Delaware, Florida, Georgia, Hawaii, Maryland Massachusetts, New York, North Carolina, Ohio, Rhode Island, Tennessee, and Washington D.C. received large grants in 2010, while Arizona, Colorado, Illinois, Kentucky, Louisiana, New Jersey, and Pennsylvania received small grants in 2011.

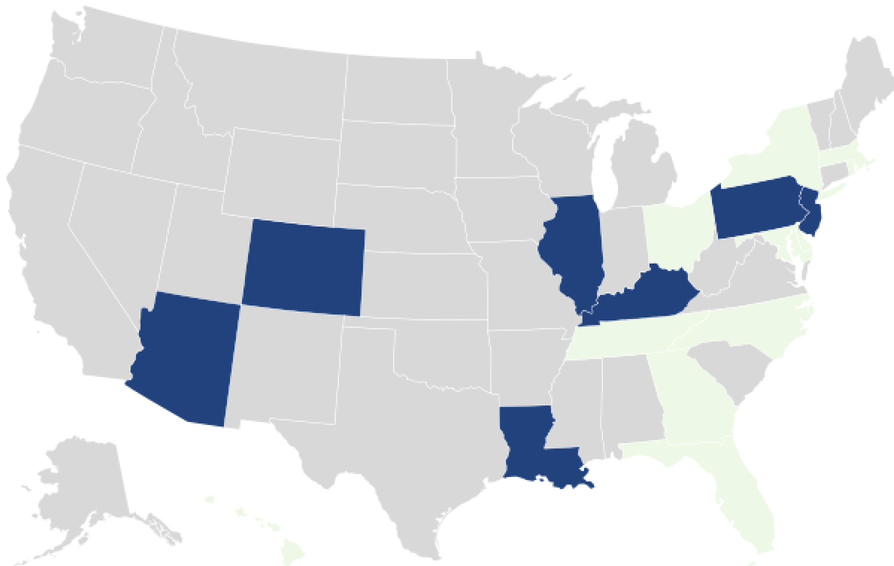


Figure 3.3: Map of Treated States

Notes: Early treated states are shown in light green, late treated states are shown in dark blue, and never treated states are shown in gray. Delaware, Florida, Georgia, Hawaii, Maryland Massachusetts, New York, North Carolina, Ohio, Rhode Island, Tennessee, and Washington D.C. received large grants in 2010, while Arizona, Colorado, Illinois, Kentucky, Louisiana, New Jersey, and Pennsylvania received small grants in 2011.

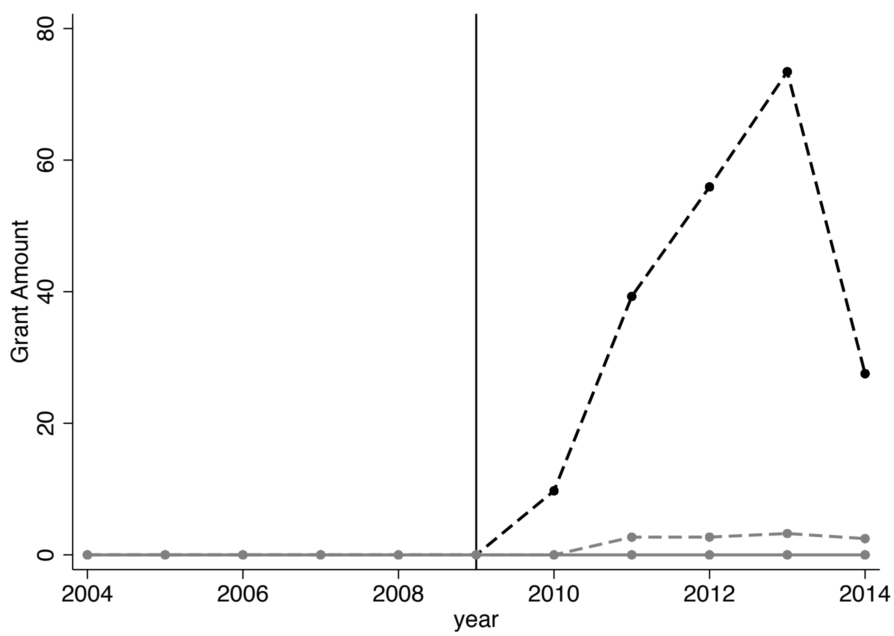


Figure 3.4: Early Treated States Received Larger Grants Than Late Treated States

Notes: This figure presents means of RTTT grants to states. The black dashed line shows average grant amount for the 2010 cohort, the gray dashed line shows average grant amount for the 2011 cohort, and the solid gray line shows means for the never treated cohort (which is mechanically zero).

Still though, late treated states remained committed to implementing RTTT policies, at least in their public statements; the same New Jersey spokesperson stated that "at the end of the day, [New Jersey stands] by these reforms and we need to move forward on them quickly" (Calefati, 2011). Similarly, a representative from the Louisiana department of education said that "the Race to the Top process has moved ... us significantly down the pathway [to reform]" despite Louisiana receiving a small grant (Paulson, 2011).

By requiring states to demonstrate the capacity to implement the reforms in their applications before announcing which states would receive grants, ED successfully bound the path of policy adoption in both early and late treated states regardless of funding. Consequently, states that received any RTTT grant money implemented more RTTT promoted policies than non-RTTT states (Boser, 2012; Howell, 2015). After grant receipt, there were no statistically distinguishable differences between early and late treated states in the number of policies meant to improve state capacity, build state data systems, recruit and retain effective teachers, turn around the lowest performing schools, or remove restrictions on charter schools (Dragoset et al., 2016).[3]

3.2 Data and Methods

Data on grants to states comes from ED and contains information on the size and timing of each state's RTTT grant. I measure state level student achievement with the percent of students scoring proficient on the National Assessment of Educational Progress, a nationally representative assessment of student achievement given to fourth and eighth graders biannually since 2003.

I examine the effect of RTTT grants on student achievement with the event study specification in equation 3.1.

$$\begin{aligned} P_{sy} = &\alpha_0 + \sum_{t=2003}^{2009} \pi_t 1\{y - 2010 = t\} * treat_s + \sum_{t=2011}^{2019} \gamma_t 1\{y - 2010 = t\} * treat_s \\ &+ \alpha_s + \alpha_{gy} + \varepsilon_{sy} \end{aligned} \tag{3.1}$$

P_{sy} is the percent of students scoring proficient on the NAEP in state s in year y, time relative to treatment is given by $1\{y - 2010 = t\}$, which is a dummy variable that equals one for observations of states t years

[3] The only difference in policy adoption between early and late treated states is related to testing; while early and late treated states adopted similar standards and wrote similar tests, early treated states implemented more policies designed to support the transition to enhanced standards and high-quality assessments. Specifically, this difference is driven by differences in whether early and late states reported (1) providing supports to districts and/or schools specifically designed to aid in the implementation of the Common Core State Standards (CCSS) with English Language Learners (ELLs) since adopting these standards, (2) investing in new technology to assist with implementation of the assessments associated with the CCSS since adopting them, and (3) providing funds to districts and/or schools to support implementation of the CCSS since adopting them (Dragoset et al., 2016). While differences in common core adoption have large effects on economically advantaged students, they have no effect on economically disadvantaged students (Bleiberg, 2020). Since the differences between early and late treated states are driven by differences in serving largely disadvantaged groups, these differences should not affect student outcomes.

before or after 2010, the year that early treated states received RTTT grants, and $treat_s$ is a dummy variable indicating states in the treatment group. α_s and α_{gy} are state and grade-by-year fixed effects.

I summarize this event study specification by estimating the following difference-in-difference:

$$P_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 2010\} + \delta_s treat_s + \alpha_{gy} + \varepsilon_{sy} \tag{3.2}$$

Cluster robust standard errors can over-reject the null hypothesis in situations where there are few treated clusters. Similarly, p-values from a wild-cluster bootstrap can either over or under-reject the null when there are few treated clusters (MacKinnon and Webb, 2018), so in addition to reporting standard errors clustered at the state level, I implement a permutation test. To do this, I construct a placebo distribution of γ from equation 3.2 by re-estimating equation 3.2 once for every possible combination of treated units.[4] I then compare my estimate of γ for the actual treatment group to the placebo distribution. In order to achieve 5 percent statistical significance using a two-tailed test, the true coefficient must be in the top (or bottom) 2.5 percentile of this distribution.[5]

Table 3.1: Summary Statistics

	(1) Never Treated	(2) 2010 Cohort	(3) 2011 Cohort
Enrollment	837,551	1,128,705	1,188,220
Percent ELL	6.6	6.3	5.6
Percent FRPL	39.2	41.6	42.4
Title I Dollars	219MM	350MM	351MM
Observations	32	12	7

Notes: This table presents the mean enrollment, percent English language learners, percent of students receiving free and reduced price lunch, and total Title I transfers to early treated states, late treated states, and never treated states in 2009.

The identifying assumption of this design is that, in absence of treatment, test scores would have followed parallel trends in early, late, and never treated states. Since RTTT grants began rolling out in 2010, lingering effects of the the Great Recession are the main threat to this assumption. Early, late, and never treated states might have been differently affected by (1) the economic downturn, causing states to alter spending or (2) by relief efforts in the ARRA besides RTTT. Since ED did not start distributing RTTT funds until 2010, the effect of either of these sources of bias should appear in the pre-treatment period of an event study. I also

[4]For example, when comparing early and late treated states, I estimate every combination of these 19 states where 12 states are treated in 2010 and 7 states are controls. In situations where there are more than 1000 combinations, I estimate a random set of 1000 combinations.

[5]Table 3.7 shows that my results are robust to performing randomization inference with cluster robust t-statistics instead, which is generally more stable than constructing p-values using coefficients (MacKinnon and Webb, 2020).

directly test for some of these sources of bias in Section 3.3.3.

To interpret comparisons of early treated states relative to late treated states as the effect of extra RTTT funding, I must assume that early and late treated states would have responded to RTTT policy changes in the same way in absence of funding.[6] Table 3.1 provides some evidence that this assumption holds by showing that early and late treated states were balanced on a number of pre-treatment observable characteristics, including enrollment, the percent of English language learners, the percent of students receiving free and reduced price lunch, and total Title I transfers.

3.3 Results

3.3.1 Effect of Funding on Test Scores

Figures 3.5 and 3.6 compare math proficiency rates between early and late treated states before and after 2010. Figure 3.5 presents estimates of the π_t and γ_t from equation 3.1 with early treated states as the treatment group and late treated states as the control group. There is no pre-treatment trend in proficiency rates, and no significant difference in proficiency after treatment. Figure 3.6 compares the corresponding difference-in-difference coefficient from equation 3.2 to the coefficient's placebo distribution. While the true coefficient is positive, it is near the middle of the placebo distribution. Randomization inference tests the sharp null that there is no effect for every treated unit as opposed to the null that the average effect is zero, so the results in Figure 3.6 do guarantee that there was no effect on test scores in every early treated state.

Figures 3.7 and 3.8 re-estiamtes the results in Figures 3.5 and 3.6 with early treated states that were in the bottom tercile of pre-treatment achievement as the treatment group and all late treated states as the control group. Figure 3.7 shows that math proficiency rates increase sharply three years after treatment in early treated states. This effect persists through the end of the study period. Figure 3.8 shows that the associated difference-in-difference coefficient is in the 99th percentile of the placebo distribution.

Table 3.2 summarizes Figures 3.5 and 3.7 by estimating equation 3.2 with math proficiency rate as the dependent variable. In column 1, the treatment group is every early treated state and the control group is every late treated state. In column 2, the treatment group is early treated states in the bottom tercile of national proficiency, and the control group is every late treated state. The effects are large and strongly significant, with math proficiency increasing by 4.3 percentage points. Columns 3 and 4 dissagregate the estimate in column 2 by grade.[7] The effects are large and significant for both groups.

[6] To identify the effect of funding exclusively, I would also have to assume that policy changes and funding are additively separable. This is unlikely to be the case, so results from this comparison should be interpreted as the effect of combining extra funding with policy changes as compared to the counterfactual of only implementing policy.

[7] See Appendix for the associated event studies.

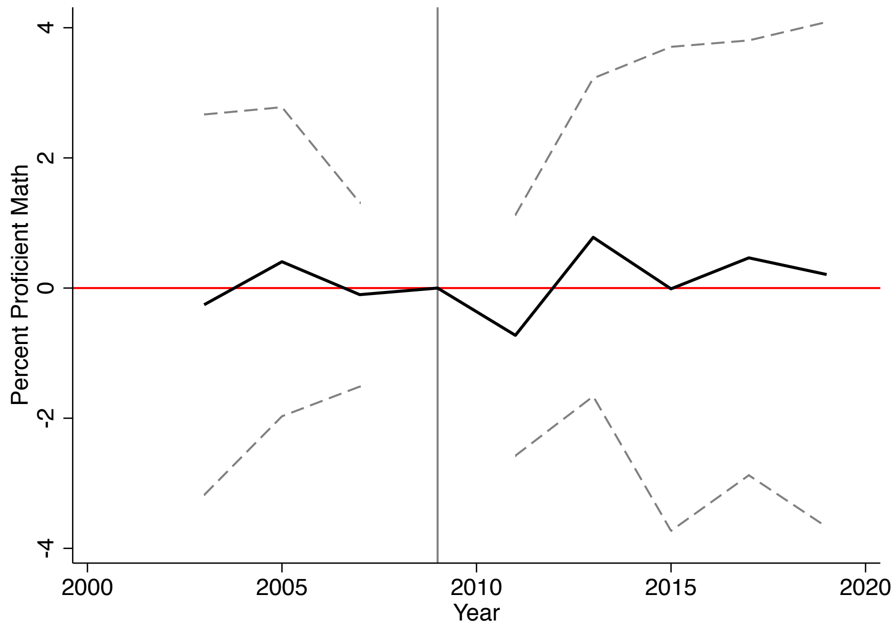


Figure 3.5: Proficiency Did Not Increase in Every Early Treated State

Notes: This figure presents estimates of the π_t and γ_t from $P_{sy} = \alpha_0 + \sum_{t=2003}^{2009} \pi_t 1\{y-2010=t\} * treat_s + \sum_{t=2011}^{2019} \pi_t 1\{y-2010=t\} * treat_s + \alpha_s + \alpha_{gy} + \varepsilon_{sy}$ where P_{sy} is state math proficiency rates on the NAEP, and $treat_s$ is a dummy variable indicating states in the 2010 cohort. α_s and α_{gy} are state and grade-by-year fixed effects. The treatment group is the 2010 cohort, and the control group is the 2011 cohort.

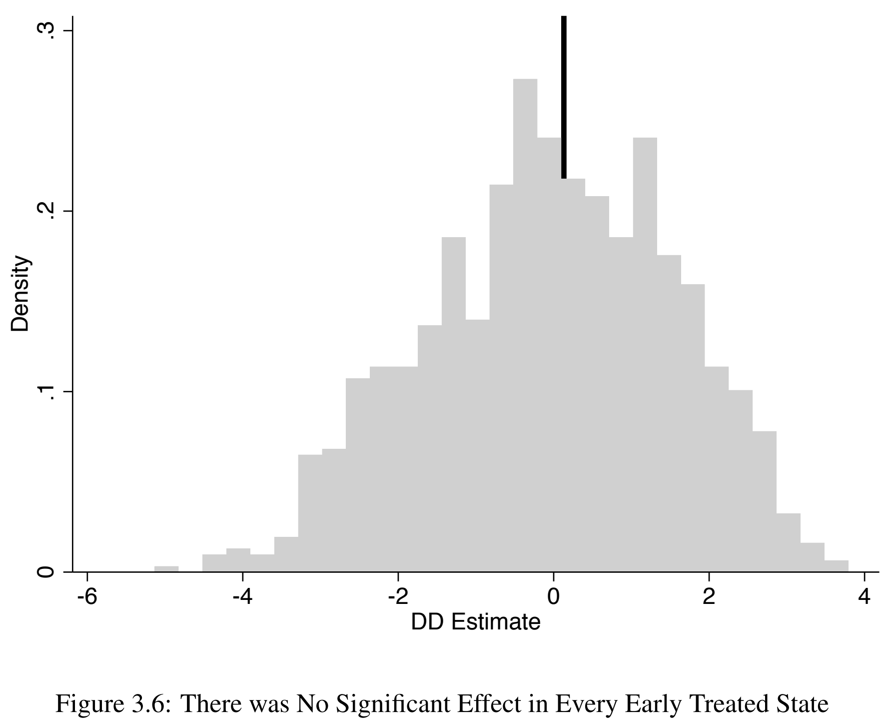


Figure 3.6: There was No Significant Effect in Every Early Treated State

Notes: This Figure presents estimates of γ from $Y_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 2010\} + \alpha_{gy} + \delta_s early_s + \varepsilon_{sy}$ compared to a placebo distribution of γ from estimating this equation on every possible treatment group. The black line indicates the true estimate of γ.

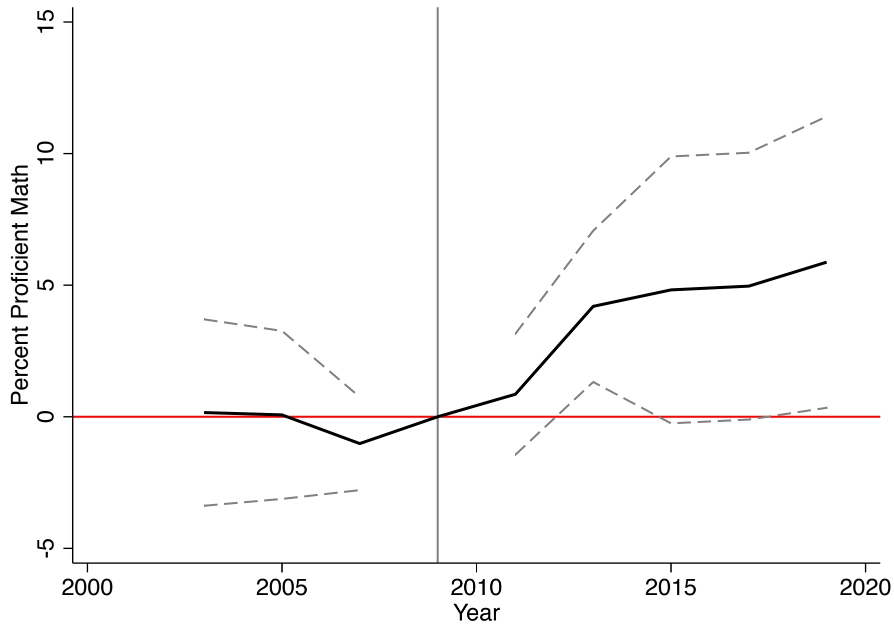


Figure 3.7: Proficiency Increased in Low Performing Early Treated States Relative to Late Treated States

Notes: This figure re-estimates the results in Figure 3.5 with states in the 2010 cohort that were in the bottom tercile of pre-treatment math proficiency as the treatment group, and the entire 2011 cohort as the control group.

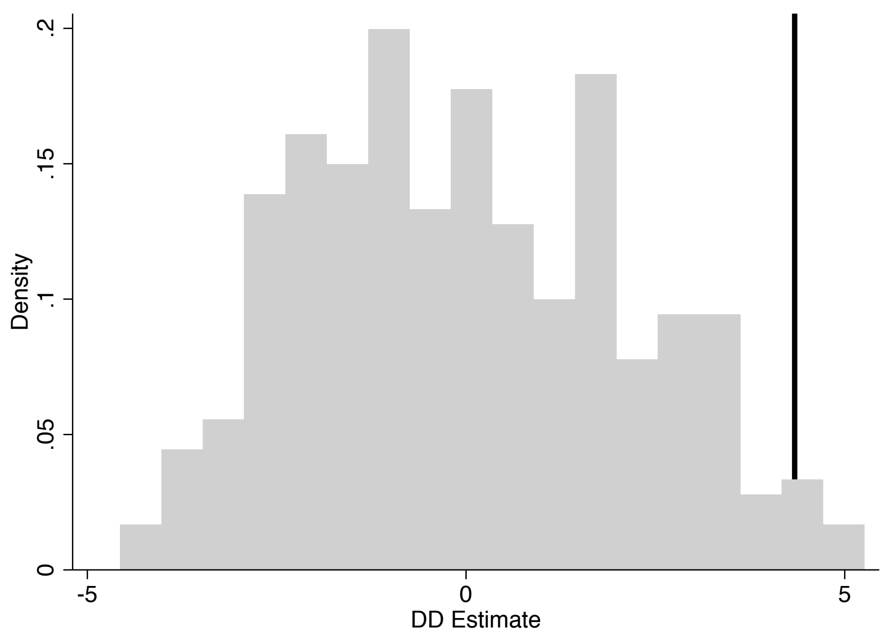


Figure 3.8: The Improvement in Low Performing Early Treated States is Significant

Notes: This figure re-estimates the results in Figure 3.6 with states in the 2010 cohort that were in the bottom tercile of pre-treatment math proficiency as the treatment group, and the entire 2011 cohort as the control group.

Column 5 shows that these results are not driven by increases in proficiency in all bottom tercile states by comparing states in the bottom tercile of achievement in the 2010 cohort to states in the bottom tercile of achievement in the 2011 cohort. While there are not enough observations to properly perform a permutation test, the results are similar in magnitude to those in column 2. Column 6 compares states in the 2010 and 2011 cohorts that were in the top two terciles of pre-treatment achievement. The effects are small and wrong-signed.

These results show that receiving a large RTTT grant led to large increases in math proficiency relative to states that received small RTTT grants, but these improvements only occurred in states that were in the bottom tercile of national achievement before treatment.

3.3.2 Effect of Policy Adoption on Test Scores

I estimate the effect of combining policy changes with new funding in Figure 3.9. This figure compares early treated states in the bottom tercile of pre-treatment achievement to never treated states in the bottom tercile of pre-treatment achievement before and after 2009.[8] Figure 3.11 re-estimates this specification with late treated states in the bottom tercile of pre-treatment achievement as the control group and never treated states in the bottom tercile of pre-treatment achievement as the control group.

Table 3.2: RTTT had Larger Effects in Early Treated States

	(1)	(2)	(3)	(4)	(5)	(6)
	Percent Proficient Math					
Treat X 1{year >= 2010}	0.131	4.339***	5.250**	3.429**	3.117	-1.056
	(1.379)	(1.989)	(2.605)	(1.650)	(2.352)	(0.876)
	((.94))	((.01))	((.016))	((.022))		((.178))
Treated Group Tercile	1,2,3	1	1	1	1	2,3
Control Group Tercile	1,2,3	1,2,3	1,2,3	1,2,3	1	2,3
Grades	4,8	4,8	4	8	4,8	4,8
Observations	342	198	99	99	126	216

Standard errors clustered at the state level in parentheses

Randomization inference p values in double parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents estimates of γ from $Y_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 2010\} + \alpha_{gy} + \delta_s early_s + \varepsilon_{sy}$ where P_{sy} is state math proficiency rates on the NAEP, and $treat_s$ is a dummy variable indicating states in the 2010 cohort. α_s and α_{gy} are state and grade-by-year fixed effects. In column 1, the treatment group is the entire 2010 cohort, and the control group is the entire 2011 cohort. In column 2, the treatment group consists of states in the 2010 cohort that were in the bottom tercile of pre-treatment math proficiency, and the entire 2011 cohort is the control group. Column 3 and 4 re-estimate column 2 for just fourth and eighth graders respectively. In column 5, the treatment group consists of states in the 2010 cohort that were in the bottom tercile of pre-treatment math proficiency, the control group consists of states in the 2011 cohort that were in the bottom tercile of pre-treatment math proficiency. In column 6, the treatment group consists of states in the 2010 cohort that were in the top two terciles of pre-treatment math proficiency, the control group consists of states in the 2011 cohort that were in the top two terciles of pre-treatment math proficiency.

[8] In this specification I compare proficiency rates before and after 2009 because some RTTT policies went into effect shortly after RTTT was announced at the beginning of 2009 rather then when ED distributed grant funds in 2010.

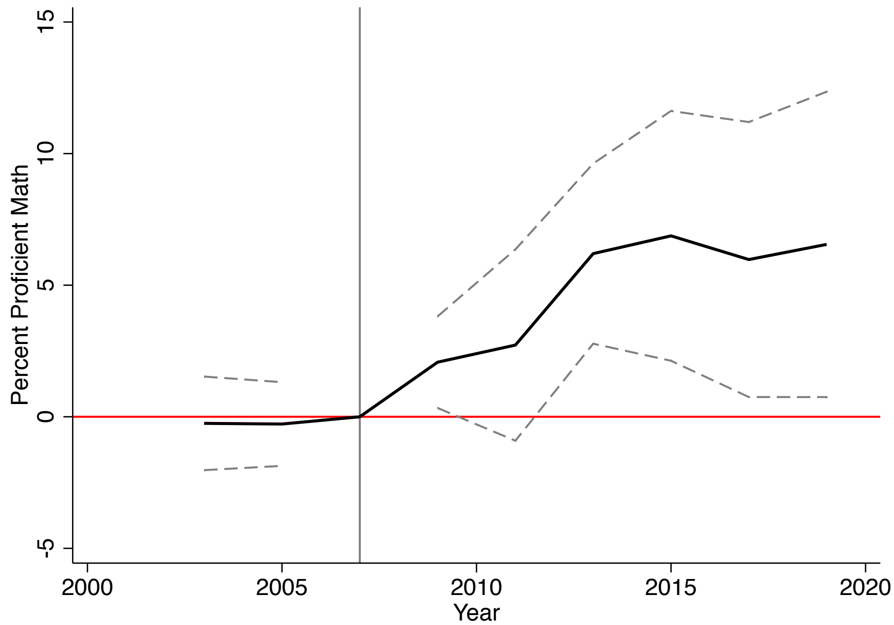


Figure 3.9: Proficiency Increased in Low Performing Early Treated States Relative to Never Treated States

Notes: This figure presents estimates of the π_t and γ_t from $Y_{sy} = \alpha_0 + \sum_{t=2003}^{2007} \pi_t 1\{y - 2009 = t\} * treat_s + \sum_{t=2009}^{2019} \pi_t 1\{y - 2009 = t\} * treat_s + \alpha_s + \alpha_{gy} + \varepsilon_{sy}$ where Y_{sy} is state math proficiency rates on the NAEP, and $treat_s$ is a dummy variable indicating states in the 2010 cohort. α_s and α_{gy} are state and grade-by-year fixed effects. The treatment group is states in the 2010 cohort in the bottom tercile of pre-treatment achievement, and the control group is never treated states in the bottom tercile of pre-treatment achievement.

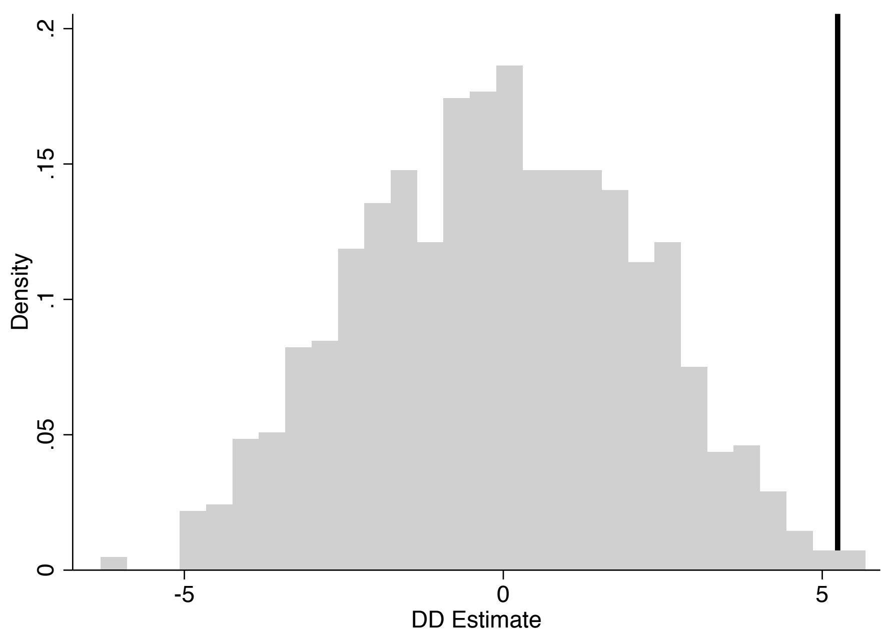


Figure 3.10: The Improvement in Low Performing Early Treated States is Significant (Never-Treated)

Notes:This figure presents estimates of γ from $Y_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 2009\} + \alpha_{gy} + \delta_s early_s + \varepsilon_{sy}$ compared to a placebo distribution of γ from estimating this equation on every possible treatment group. The black line indicates the true estimate of γ.

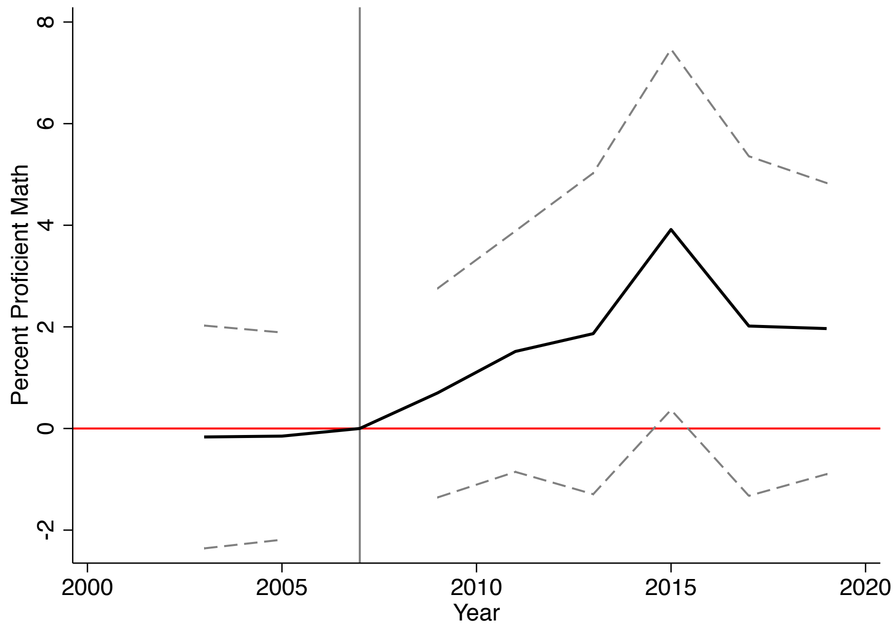


Figure 3.11: Math Proficiency Increased in Low Performing Late Treated States Relative to Never Treated States

Notes: This figure re-estimates the results in Figure 3.9 with states in the 2011 cohort in the bottom tercile of pre-treatment achievement as the treatment group.

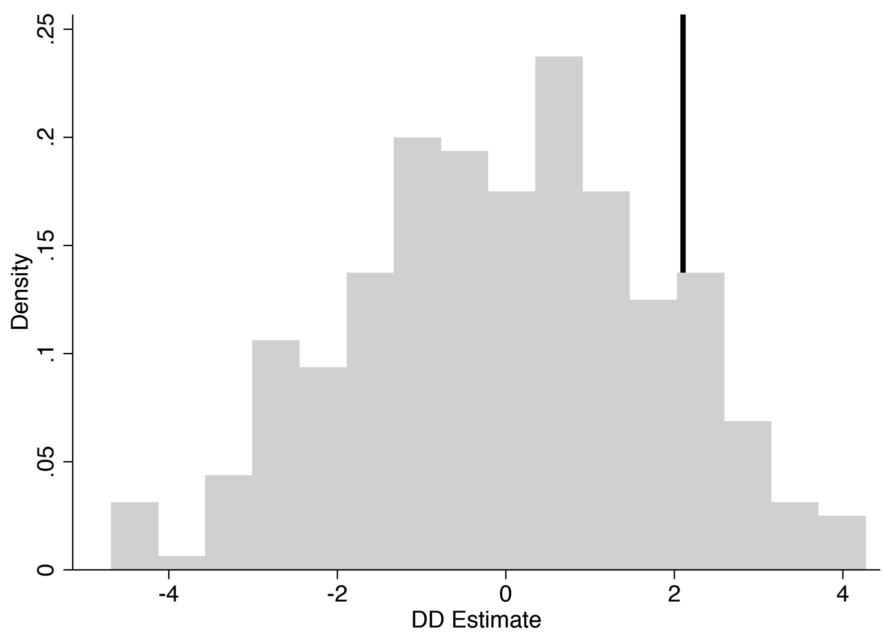


Figure 3.12: The Improvement in Low Performing Late Treated States is Marginally Significant

Notes: This figure re-estimates the results in Figure 3.10 with states in the 2011 cohort in the bottom tercile of pre-treatment achievement as the treatment group.

Table 3.3: RTTT Policies Led to Increased Test Scores in States with Low Baseline Achievement

	(1)	(2)
	2010 vs Never	2011 vs Never
	Percent Proficient Math	
Treat X 1{year >= 2009}	5.242**	2.103*
	(1.925)	(1.639)
	((.024))	((.078))
Observations	252	234

Standard errors clustered at the state level in parentheses

Randomization inference p values in double parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table estimates $Y_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 209\} + \alpha_{gy} + \delta_s treat_s + \varepsilon_{sy}$ where $treat_s$ indicates RTTT treated and the $alpha_{gy}$ are grade-by-year fixed effects. The dependent variable is the percent of students scoring proficient on the NAEP math test. Column 1 compares compares states in the 2010 cohort in the bottom tercile of pre-RTTT achievement to never treated states in the bottom tercile of pre-RTTT achievement, and column 2 compares compares states in the 2011 cohort in the bottom tercile of pre-RTTT achievement to never treated states in the bottom tercile of pre-RTTT achievement.

Table 3.3 summarizes Figures 3.9 and 3.11. Column 1 shows that proficiency rates improved by 5.2 percentage points in early treated states in the bottom tercile of achievement relative to never treated states in the bottom tercile of achievement. Column 2 shows that these effects were substantially smaller in late treated states in the bottom tercile of achievement with an estimated increase of 2.1 percentage points that is significant at the ten percent level.

3.3.3 Robustness Checks

3.3.3.1 Incorporating Data from Before 2003

While there is no evidence of a pre-trend in either Figure 3.9 or Figure 3.11, there are only three pre-treatment periods in either specification. To address this, I use additional data on NAEP scores from 1992, 1996 and 2000 to extend the pre-treatment period in Figures 3.13 and 3.14. The NAEP was not administered in every state in every year before 2003, so these event study coefficients are not balanced. Consequently, the estimates are imprecise, however, they do not reveal any pre-treatment trend in math proficiency rates.

3.3.3.2 Other Transfers

The largest federal transfer program to schools is Title I. ED allocates Title I dollars based on the number of economically disadvantaged students in a school (Stephenson and Kaiser, 2018), so Title I allocations will increase when the economic conditions in a state worsen. Table 3.4 test if these changes are correlated with treatment by estimating equation 3.2 with per-student Title I dollars as the dependent variable. In column 1, the 2010 cohort is the treated group and never treated states are the control group; in column 2, the 2011 cohort is the treated group and never treated states are the control group; in column 3, the 2010 cohort is the treated group and the 2011 cohort is the control group. This shows that regular Title I transfers remained similar across early treated states and never treated states, late treated states and never treated states, and early treated states and late treated states.

Table 3.4: Changes in Title I Were Not Correlated with Treatment

	(1) 2010 vs Never	(2) 2011 vs Never	(3) 2010 vs 2011
		Title I Dollars	
Treat X 1{year >= 2010}	-3.641	0.300	-7.258
	(10.15)	(7.583)	(11.66)
	((.568))	((.948))	((.594))
Observations	484	429	209

Standard errors clustered at the state level in parentheses
Randomization inference p values in double parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table estimates $Y_{sy} = \alpha_0 + \gamma treat_s * 1\{y >= 2010\} + \delta_y 1\{y >= 2010\} + \delta_s treat_s + \varepsilon_{sy}$. The dependent variable is average Title I dollars. $treat_s$ is a dummy variable indicating the treatment group: In column 1, early treated states are the treatment group and never treated states are the control group; in column 2, late treated states are the treatment group and never treated states are the control group; in column 3, early treated states are the treatment group and late treated states are the control group.

Along with RTTT grants, the ARRA distributed extra money to districts according to existing Title I formulas, as well as extra education funding to states through the State Fiscal Stabilization Fund (SFSF) in 2009. Table 3.5 shows that, in per student terms, early treated states received more ARRA and SFSF dollars, though these differences were much smaller than the differences in RTTT grant size. While, to the best of my knowledge, there has not been a causal study of the effect of ARRA education transfers on spending, accounting studies suggest that states generally used these transfers to plug budget holes rather than increase spending (Superfine, 2011), and I provide evidence that these transfers did not cause changes in local education spending in the Appendix.

Table 3.5: The 2010 Cohort Received More ARRA Title I Dollars

	(1) Never Treated	(2) 2010 Cohort	(3) 2011 Cohort
Recovery Title I Amount	13.6MM	22.6MM	21.1MM
SFSF Amount	793MM	1,160MM	1200MM
RTTT Amount	0	328MM	28.5MM
Recovery Title I Amount	187.71	238.23	191.83
SFSF Amount	960.90	1051.34	976.69
RTTT Amount	0	435.64	23.68
Observations	32	12	7

Notes: This table presents the mean Title I transfers from ARRA, SFSF transfers, and RTTT transfers to early treated states, late treated states, and never treated states. The top panel presents average totals, and the bottom panel presents average per student transfers.

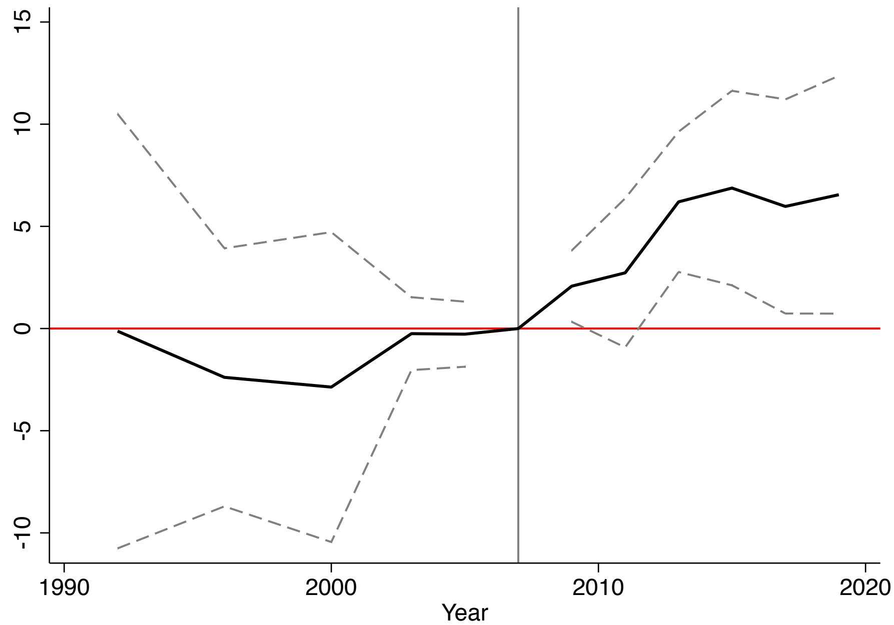


Figure 3.13: There was no Pre-Trend in Early Treated States

Notes: This Figure re-estimates the results in Figure ?? incorporating unbalanced data on NAEP scores from 1992, 1996, and 2000.

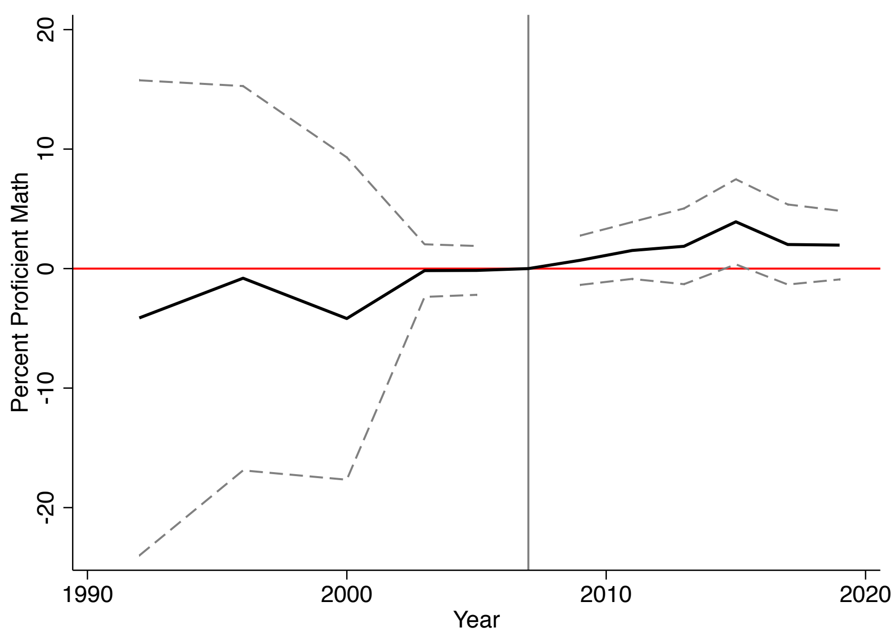


Figure 3.14: There was no Pre-Trend in Late Treated States

Notes: This Figure re-estimates the results in Figure ?? incorporating unbalanced data on NAEP scores from 1992, 1996, and 2000.

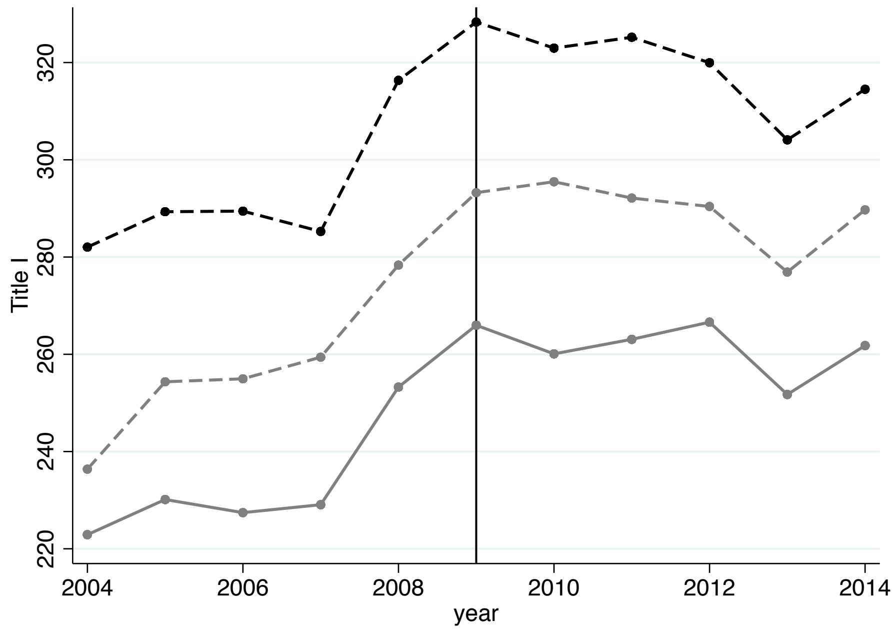


Figure 3.15: Title I Dollars by Treatment

Notes: This figure presents means of total Title I transfers to districts in early, late, and never treated states. The black dashed line shows per student Title I dollars for the 2010 cohort, the gray dashed line shows per student Title I dollars for the 2011 cohort, and the solid gray line shows per student Title I dollars for the never treated cohort.

3.4 Conclusion

By successfully encouraging policy changes that led to improvements in student achievement in states that received both large and small grants, RTTT demonstrates the ability of competitive grant programs to improve second stage outcomes at relatively low costs. The improvements in achievement were larger in early treated states, which suggests that, while competitive grant programs can have a meaningful effect even in places that do not win grants, policy is more effective when it is paired with extra funding.

I conclude by noting two potential drawbacks of competitive grant programs: (1) policymakers likely pay a political cost for making states compete for grant funding, and (2) the states most likely to win grants are those that have the time and funding available to construct compelling grant applications. For this reason, the states most likely to win competitive grant funding may be those that need it the least.

3.5 Appendix

3.5.1 The Effect of Stimulus Money on Spending

Table 3.5 shows that, in per student terms, early treated states received more ARRA and SFSF dollars than late treated states and never treated states, though these differences were much smaller than the differences in RTTT grant size. While there has not been a causal study of the effect of ARRA education transfers on spending, descriptive studies suggest that these transfers were generally used to plug budget holes rather than increase spending (Superfine, 2011). In this section, I provide additional evidence that these transfers did not increase spending using district level spending data from the Census Bureau's Annual Survey of School System Finances.

ED distributed extra education funding from ARRA according to existing Title I formulas. Title I funding is distributed to schools according to the number of disadvantaged students that they serve, so I can predict how much ARRA funding a district will receive by calculating the percent of the age 5-17 population that was living in poverty in a districts catchment area in 2009. I do this using the Census Bureau's Small Area Income and Poverty Estimates (SAIPE), which produces single-year estimates of the number of school-age children living in poverty for all school districts in the US. This is the same data that ED uses when calculating Title I allocations (Stephenson and Kaiser, 2018).

I can estimate the effect of extra education funding from ARRA by leveraging continuous variation in school age poverty using equation 3.3,

$$Y_{dy} = \alpha_0 + \sum_{t=2004}^{2007} \pi_t 1\{y-2009=t\} * pct_{dy} + \sum_{t=2009}^{2010} \gamma_t 1\{y-2009=t\} * pct_{dy} + \alpha_d + \alpha_y + \varepsilon_{dy} \tag{3.3}$$

where d indexes districts, y indexes years, and pct_{dy} is the percent of the age 5-17 population living within a districts service area that is living in poverty.

I can summarize this effect with equation 3.4.

$$Y_{dy} = \alpha_0 + \gamma pct_{dy} + \alpha_d + \alpha_y + \varepsilon_{dy} \tag{3.4}$$

Figure 3.16 shows that pct_{dy} accurately predicts federal revenue by estimating equation 3.3 with federal revenue as the dependent variable. Figure 3.17 then re-estimates this specification with total expenditure as the dependent variable.

Table 3.6 summarizes these figures. While federal dollars increased, this did not lead to an equivalent increase in local expenditure, suggesting that receiving school districts used extra funding to offset budget shortfalls caused by the Great Recession rather than increasing total spending.

3.5.2 Additional Figures and Tables

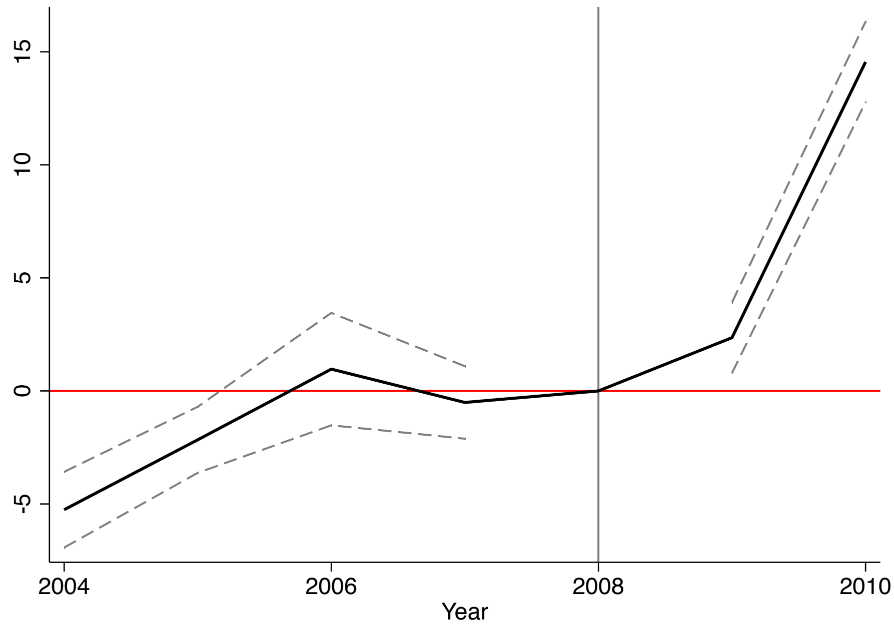


Figure 3.16: Federal Revenue Increased in Districts with a Higher Percentage of Disadvantaged Students

Notes: This figure presents estimates of the γ_t and π_t from $Y_{dy} = \alpha_0 + \sum_{t=2004}^{2007} \pi_t 1\{y - 2009 = t\} * pct_{dy} + \sum_{t=2009}^{2010} \gamma_t 1\{y - 2009 = t\} * pct_{dy} + \alpha_d + \alpha_y + \varepsilon_{dy}$. The dependent variable is total federal revenue.

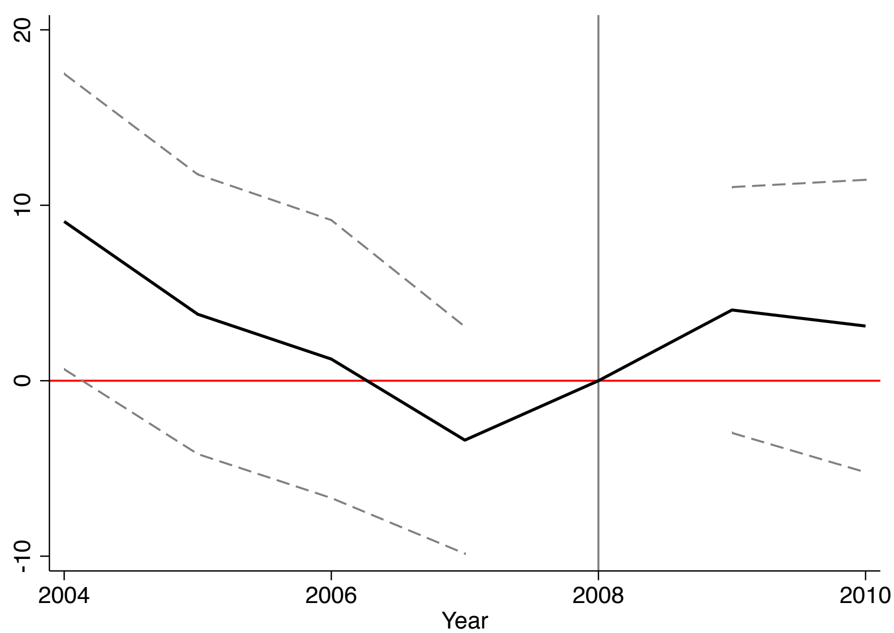


Figure 3.17: Total Expenditure did not Change in Districts with More Disadvantaged Students

Notes: This Figure re-estimates the the specification in Figure 3.16 with total expenditure by district d in year y as the dependent variable.

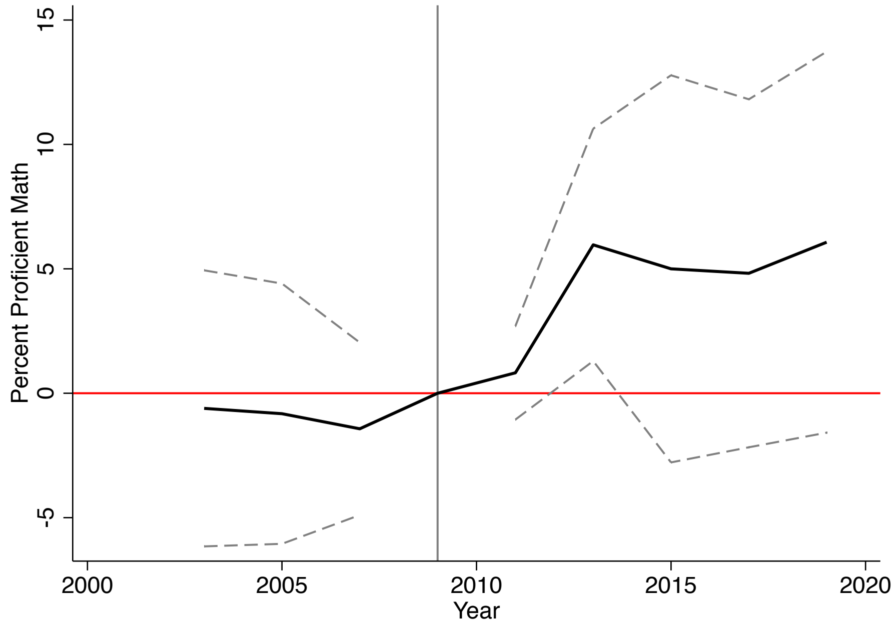


Figure 3.18: Fourth Grade Percent Proficient Event Study (Bottom Tercile)

This figure re-estimates the results in Figure 3.7 on fourth grade test scores.

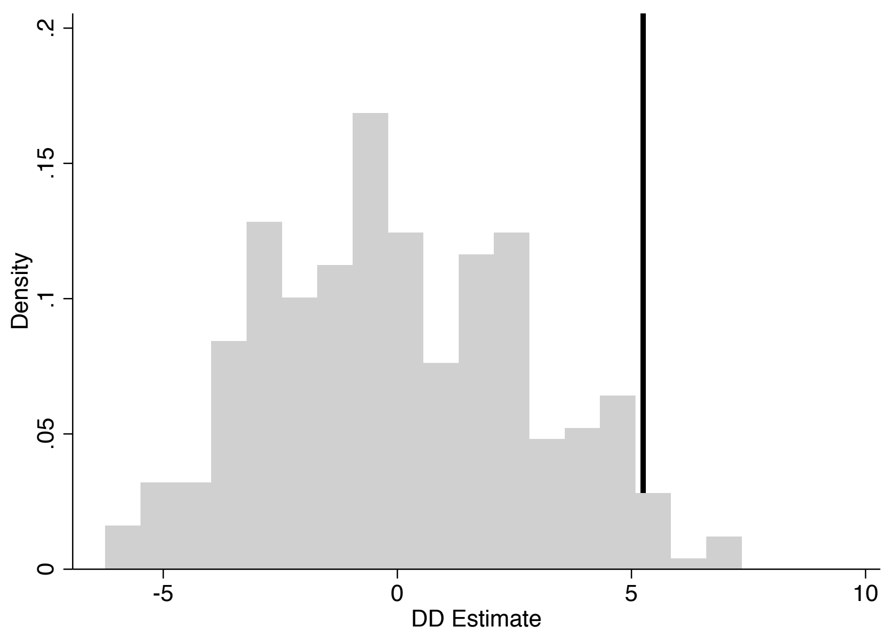


Figure 3.19: Fourth Grade Percent Proficient Difference-in-Differences (Bottom Tercile)

This figure re-estimates the results in Figure 3.7 on fourth grade test scores.

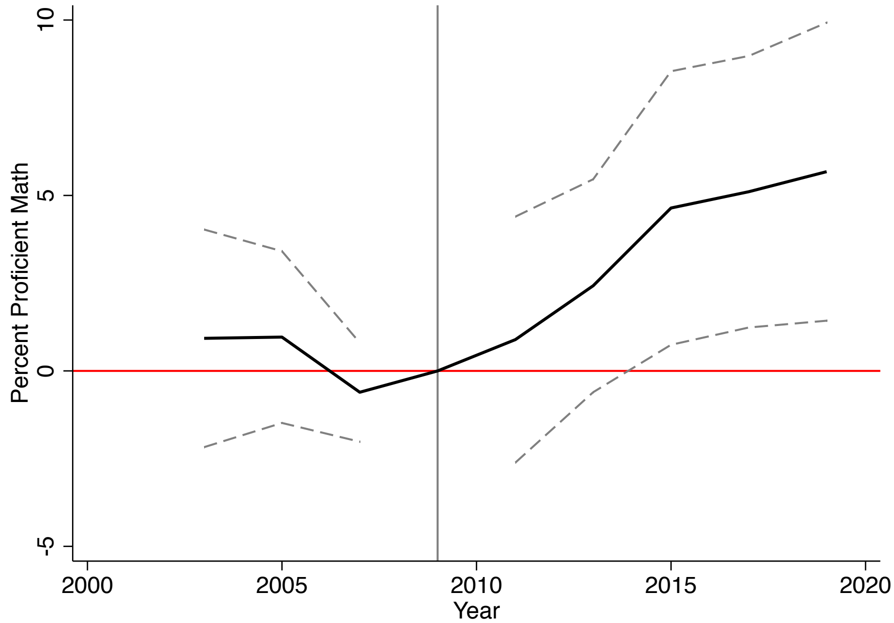


Figure 3.20: Eighth Grade Percent Proficient Event Study (Bottom Tercile)

This figure re-estimates the results in Figure 3.7 on eighth grade test scores.

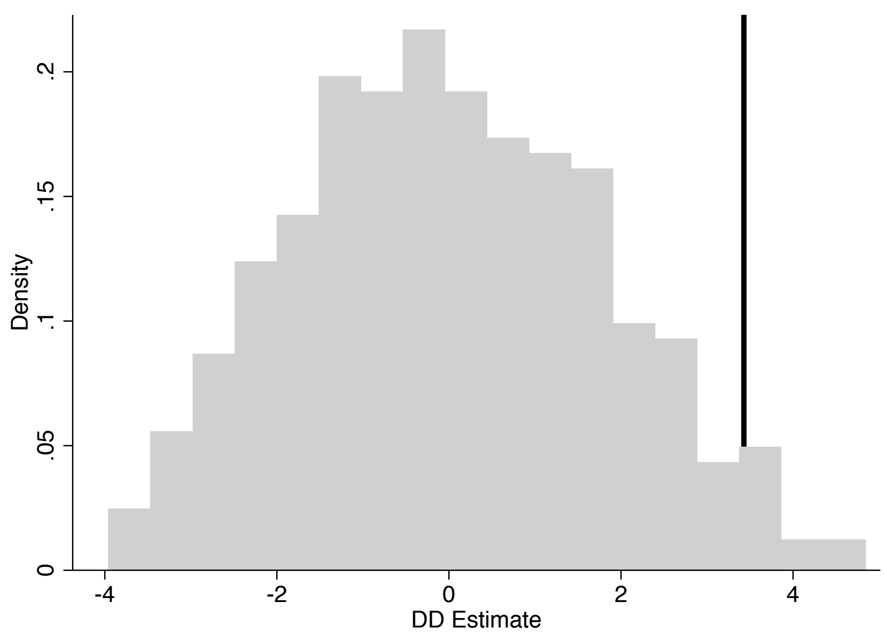


Figure 3.21: Eighth Grade Percent Proficient Difference-in-Differences (Bottom Tercile)

This figure re-estimates the results in Figure 3.7 on eighth grade test scores.

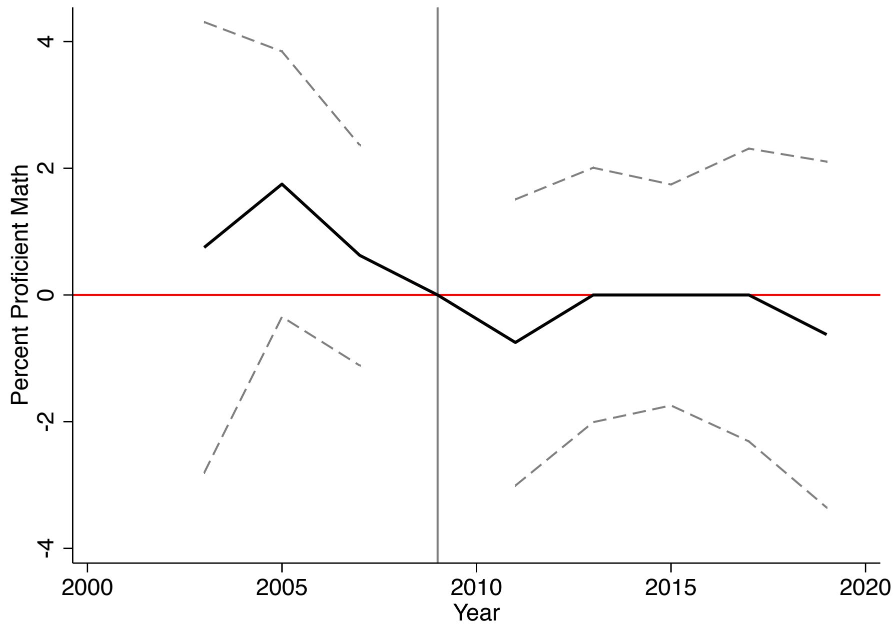


Figure 3.22: Percent Proficient Event Study (Top Two Terciles)

This figure re-estimates the results in Figure 3.5 with states in the 2010 cohort that were in the top two terciles of pre-treatment math proficiency as the treatment group, and states in the 2011 cohort that were in the top two terciles of pre-treatment math proficiency as the control group.

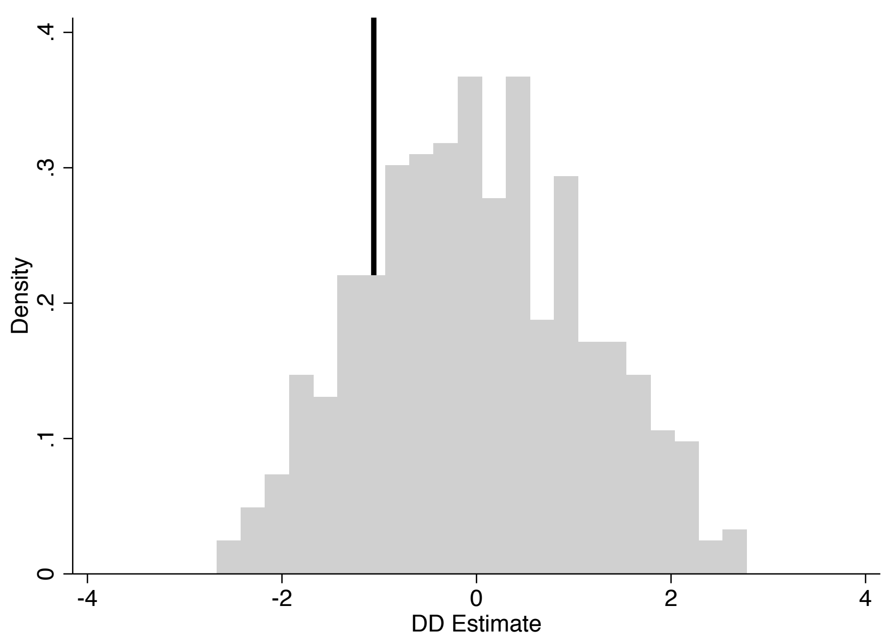


Figure 3.23: Percent Proficient Difference-in-Differences (Top Two Terciles)

This figure re-estimates the results in Figure 3.5 with states in the 2010 cohort that were in the top two terciles of pre-treatment math proficiency as the treatment group, and states in the 2011 cohort that were in the top two terciles of pre-treatment math proficiency as the control group.

Table 3.6: The Effect of Stimulus Money on District Spending

	(1) Total Federal Revenue	(2) Total Expenditure
Percent Poverty X 1{year >= 2009}	9.847***	1.431
	(0.754)	(3.122)
Observations	90107	90107

Standard errors clustered at the district level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

This table presents estimates of $Y_{dy} = \alpha_0 + \gamma pct_{dy} + \alpha_d + \alpha_y + \varepsilon_{dy}$ where pct_{dy} is the percent of the age 5-17 population living within district d's service area living in poverty in fiscal year 2009, and α_d and α_y are district and year fixed effects. The dependent variable is total federal revenue in column 1, and total expenditure in column 2.

Table 3.7: Test Scores Results (t-statistic)

	(1) 2010 vs Never	(2) 2011 vs Never	(3) 2010 vs 2011
	Percent Proficient Math		
Treat X Post	5.242**	2.103*	4.339***
	(1.925)	(1.639)	(1.989)
	((.028))	((.09))	((.01))
Observations	252	234	198

Standard errors clustered at the state level in parentheses

Randomization inference p values in double parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table re-estimates the main results from Tables 3.2 and comparing the t statistic to a placebo distribution of t statistics instead of comparing to a placebo distribution of difference-in-difference coefficients. Column 1 presents the result from column 2 of Table 3.2, and columns 2 and 3 presents results from columns 1 and 2 of Table 3.3.

References

Abdel-Raouf, N., Al-Homaidan, A., and Ibraheem, I. (2012). Microalgae and wastewater treatment. *Saudi journal of biological sciences*, 19(3):257–275.

Almond, D., Chay, K. Y., and Lee, D. S. (2005). The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083.

Almond, D. and Mazumder, B. (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics*, 3(4):56–85.

Alsan, M. and Goldin, C. (2019). Watersheds in child mortality: the role of effective water and sewerage infrastructure, 1880–1920. *Journal of Political Economy*, 127(2):586–638.

Anderson, D. M., Charles, K. K., and Rees, D. I. (2020). Re-examining the contribution of public health efforts to the decline in urban mortality. *American Economic Journal: Applied Economics*.

Arizona Department of Water Resources (2020). Cws service area. Technical report.

Arkansas GIS Office (2013). Public water systems. Technical report.

Baicker, K. (2001). Government decision-making and the incidence of federal mandates. *Journal of Public Economics*, 82:147–194.

Baicker, K. and Gordon, N. (2006). The effect of state education finance reform on total local resources. *Journal of Public Economics*, 90:1519–1535.

Beach, B., Ferrie, J., Saavedra, M., and Troesken, W. (2016). Typhoid fever, water quality, and human capital formation. *The Journal of Economic History*, 76(1):41–75.

Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and statistics*, 86(2):586–601.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439.

Bleiberg, J. (2020). Does the common core have a common effect?: An exploration of effects on academically vulnerable students. (20-213).

Boser, U. (2012). Race to the top: What have we learned from the states so far? a state-by-state evaluation of race to the top performance. *Center for American Progress*.

Brunner, E., Hyman, J., and Ju, A. (2020). School finance reforms, teachers' unions, and the allocation of school resources. *Review of Economics and Statistics*, 102(3):473–489.

Bureau of Economic Analysis (2012-2017). State outdoor recreation data.

Calefati, J. (2011). N.j. receives $38 million in race to the top education funding. *The Star Ledger*.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. C. (2021). Difference-in-differences with a continuous treatment.

Callaway, B. and Sant'Anna, P. H. (2020a). did: Difference in differences. R package version 2.0.1.906.

Callaway, B. and Sant'Anna, P. H. (2020b). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Camacho, A. (2008). Stress and birth weight: evidence from terrorist attacks. *American Economic Review*, 98(2):511–15.

Card, D. and Payne, A. A. (2002). School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics*, 83(1):49 – 82.

Chahal, C., Van Den Akker, B., Young, F., Franco, C., Blackbeard, J., and Monis, P. (2016). Pathogen and particle associations in wastewater: significance and implications for treatment and disinfection processes. *Advances in applied microbiology*, 97:63–119.

Cheng, A. S. and Dale, L. (2020). Achieving adaptive governance of forest wildfire risk using competitive grants: Insights from the colorado wildfire risk reduction grant program. *Review of Policy Research*, 37(5):657–686.

Connecticut State Department of Public Health (2020). Buffered community public water supply service areas. Technical report.

Copeland, C. (2016). Clean water act: a summary of the law. Technical report, Congressional Research Service.

Courant, P., Gramlich, E., and Rubinfeld, D. (1979). The stimulative effects of intergovernmental grants or why money sticks where it hits, in "fiscal federalism and grants-in-aid"(mieszkowski, p. and oakland, wh, eds.). *Coupe Papers on Public Economics, Urban Institute, Washington DC*.

Craig, S. and Inman, R. P. (1986). Education, welfare and the" new" federalism: State budgeting in a federalist public economy. In *Studies in state and local public finance*, pages 187–228. University of Chicago Press.

Currie, J., Graff Zivin, J., Meckel, K., Neidell, M., and Schlenker, W. (2013). Something in the water: Contaminated drinking water and infant health. *Canadian Journal of Economics/Revue canadienne d'économique*, 46(3):791–810.

Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century united states. *Demography*, 42(1):1–22.

Daigger, G. T. (1998). *Upgrading wastewater treatment plants*, volume 2. CRC Press.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–48.

Dieter, C. A. (2018). *Water availability and use science program: Estimated use of water in the United States in 2015*. Geological Survey.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., Boyle, A., Tanenbaum, C., Giffin, J., and Upton, R. (2016). Race to the top: Implementation and relationship to student outcomes. Technical report, Mathematica Policy Research.

Ferrie, J. P. and Troesken, W. (2008). Water and chicago's mortality transition, 1850–1925. *Explorations in Economic History*, 45(1):1–16.

Figlio, D., Guryan, J., Karbownik, K., and Roth, J. (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104(12):3921–55.

Fisher, R. C. and Papke, L. E. (2000). Local government responses to education grants. *National Tax Journal*, 53(1):153–168.

Flynn, P. and Marcus, M. M. (2021). A watershed moment: The clean water act and infant health. Working Paper 29152, National Bureau of Economic Research.

Flynn, P. and Smith, T. (2021). Rivers, lakes and revenue streams: The heterogeneous effects of clean water act grants on local spending. Technical report, Working Paper at Vanderbilt University Department of Economics.

Forstall, R. (1995). Population of counties by decennial census: 1900 to 1990. Technical report.

Freeman, A. M. (2010). Water pollution policy. In *Public policies for environmental protection*, pages 179–224. Routledge.

Gamkhar, S. (2000). Is the response of state and local highway spending symmetric to increases and decreases in federal highway grants? *Public Finance Review*, 28(1):3–25.

Gamkhar, S. (2003). The role of federal budget and trust fund institutions in measuring the effect of federal highway grants on state and local government highway expenditure. *Public Budgeting & Finance*, 23(1):1–21.

Gamkhar, S. and Oates, W. (1996). Asymmetries in the response to increases and decreases in intergovernmental grants: Some empirical findings. *National tax journal*, pages 501–512.

Goodman-Bacon, A. (2021a). Difference-in-differences with variation in treatment timing. Technical report.

Goodman-Bacon, A. (2021b). Drdecomp: Stata module to perform a bacon decomposition of dose-response difference-in-differences estimation.

Goodman-Bacon, A., Goldring, T., and Nichols, A. (2019). Bacondecomp: Stata module to perform a bacon decomposition of difference-in-differences estimation.

Gordon, N. (2004). Do federal grants boost school spending? evidence from title i. *Journal of Public Economics*, 88(9-10):1771–1792.

Gramlich, E. M. (1969). State and local governments and their budget constraint. *International Economic Review*, 10(2):163–182.

Grossman, D. S. and Slusky, D. J. (2019). The impact of the flint water crisis on fertility. *Demography*, 56(6):2005–2031.

Grossman, J. B. and Roberts, J. (1989). Welfare savings from employment and training programs for welfare recipients. *The Review of Economics and Statistics*, pages 532–537.

Helm, I. and Stuhler, J. (2021). The dynamic response of municipal budgets to revenue shocks.

Henderson, J. M. (1968). Local government expenditures: A social welfare analysis. *The Review of Economics of Statistics*, pages 156–163.

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

Henry, G. T., Guthrie, J. E., and Townsend, L. (2015). Outcomes and impacts of north carolina's initiative to turn around the lowest-achieving schools. *The Friday Institute for Educational Innovation, North Carolina State University. Google Scholar*.

Hines, J. R. and Thaler, R. H. (1995). The flypaper effect. *Journal of economic perspectives*, 9(4):217–226.

Hines, N. W. (1966). Nor any drop to drink: Public regulation of water quality part i: State pollution control programs. *Iowa L. Rev.*, 52:186.

Howell, W. G. (2015). Results of president obama's race to the top. *Education Next*, 15(4):58–67.

Huffman, W. E. and Evenson, R. E. (2006). Do formula or competitive grant funds have greater impacts on state agricultural productivity? *American Journal of Agricultural Economics*, 88(4):783–798.

Inman, R. P. (2008). The flypaper effect. Working Paper 14579, National Bureau of Economic Research.

Jerch, R. L. (2018). The local consequences of federal mandates: Evidence from the clean water act. Technical report, Working paper, Johns Hopkins University.

Kansas Water Office (2020). Public water supply system. Technical report.

Keiser, D. A., Kling, C. L., and Shapiro, J. S. (2019). The low but uncertain measured benefits of us water quality policy. *Proceedings of the National Academy of Sciences*, 116(12):5262–5269.

Keiser, D. A. and Shapiro, J. S. (2019a). Consequences of the clean water act and the demand for water quality. *Quarterly Journal of Economics*.

Keiser, D. A. and Shapiro, J. S. (2019b). Us water pollution regulation over the past half century: Burning waters to crystal springs? *Journal of Economic Perspectives*, 33(4):51–75.

Kniesner, T. J. and Viscusi, W. K. (2019). The value of a statistical life. *Forthcoming, Oxford Research Encyclopedia of Economics and Finance*, pages 19–15.

Knight, B. (2002). Endogenous federal grants and crowd-out of state government spending: Theory and evidence from the federal highway aid program. *American Economic Review*, 92(1):71–92.

Kuwayama, Y., Olmstead, S., and Zheng, J. (2018). The value of water quality: Separating amenity and recreational benefits. Technical report, Working Paper at the University of Texas, Austin LBJ School of Public Policy.

Leduc, S. and Wilson, D. (2017). Are state governments roadblocks to federal stimulus? evidence on the flypaper effect of highway grants in the 2009 recovery act. *American Economic Journal: Economic Policy*, 9(2):253–92.

Los Angeles Times (2010). California rejoins race to the top race, sort of.

Lutz, B. (2010). Taxation with representation: Intergovernmental grants in a plebiscite democracy. *The Review of Economics and Statistics*, 92(2):316–332.

Lutz, B. (2015). Quasi-experimental evidence on the connection between property taxes and residential capital investment. *American Economic Journal: Economic Policy*, 7(1):300–330.

Lyon, R. M. and Farrow, S. (1995). An economic analysis of clean water act issues. *Water Resources Research*, 31(1):213–223.

MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal*, 21(2):114–135.

MacKinnon, J. G. and Webb, M. D. (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2):435–450.

Marcus, M. (2021). Testing the water: Drinking water quality, public notification, and child outcomes. *The Review of Economics and Statistics*, pages 1–45.

McGuire, M. (1978). A method for estimating the effect of a subsidy on the receiver's resource constraint: with an application to us local governments 1964–1971. *Journal of Public Economics*, 10(1):25–44.

Megdal, S. B. (1987). The flypaper effect revisited: An econometric explanation. *The Review of Economics and Statistics*, pages 347–351.

Mello, S. (2019). More cops, less crime. *Journal of Public Economics*, 172:174–200.

National Center for Health Statistics (1968-1988a). Natality detail data.

National Center for Health Statistics (1968-1988b). Vital statistics compressed mortality file.

National Environmental Research Center (1972). Upgrading existing wastewater treatment plants. Technical report.

New Jersey Department of Environmental Protection (2004). Njdep public community water purveyor service areas, 1998. Technical report.

North Carolina Department of Environmental Quality (2017). Public water supply water sources. Technical report.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of economic literature*, 37(3):1120–1149.

Oreopoulos, P., Stabile, M., Walld, R., and Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. *Journal of human Resources*, 43(1):88–138.

Paulson, A. (2011). Race to the top losers: Why did louisiana and colorado fail? *Yahoo News*.

Pennsylvania Spatial Data Access (2015). Public water supply 2015. Technical report.

Peterson, W. and Rothstein, R. (2010). Let's do the numbers: Department of education's 'race to the top'program offers only a muddled path to the finish line. *Economic Policy Institute Briefing Paper*, 263.

Raff, Z. and Earnhart, D. (2019). The effects of clean water act enforcement on environmental employment. *Resource and Energy Economics*, 57:1–17.

Rassier, D. G. and Earnhart, D. (2010). The effect of clean water regulation on profitability: testing the porter hypothesis. *Land Economics*, 86(2):329–344.

Reynolds, K. A., Mena, K. D., and Gerba, C. P. (2008). Risk of waterborne illness via drinking water in the united states. In *Reviews of environmental contamination and toxicology*, pages 117–158. Springer.

Roth, J. and Sant'Anna, P. H. (2021). Efficient estimation for staggered rollout designs. *arXiv preprint arXiv:2102.01291*.

Royer, H. (2009). Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85.

Rubin, K. (1985). Efficient investments in wastewater treatment plants. US Congress, Congressional Budget Office.

Rundquist, J. and Calefati, J. (2010). Error on 'race to the top' application costs n.j. $400m in federal funds. *The Star Ledger*.

Solley, W. B., Merk, C. F., and Pierce, R. R. (1988). Estimated use of water in the united states in 1985. Technical report.

Stephenson, T. and Kaiser, P. (2018). Title i allocation formulas. Technical report.

Strumpf, K. S. (1998). A predictive index for the flypaper effect. *Journal of Public Economics*, 69(3):389–412.

Superfine, B. M. (2011). Stimulating school reform: The american recovery and reinvestment act and the shifting federal role in education. *Mo. L. Rev.*, 76:81.

Texas Commission on Environmental Quality (2020). Public drinking water supply. Technical report.

Toccalino, P. and Hopple, J. A. (2010). *The Quality of Our Nation's Waters: Quality of Water from Public-supply Wells in the United States, 1993-2007: Overview of Major Findings*. US Geological Survey.

Travis, R., Morris, J. C., and Morris, E. D. (2004). State implementation of federal environmental policy: Explaining leveraging in the clean water state revolving fund. *Policy Studies Journal*, 32(3):461–480.

Troesken, W. (2001). Race, disease, and the provision of water in american cities, 1889–1921. *The Journal of Economic History*, 61(3):750–776.

Troesken, W. (2002). The limits of jim crow: Race and the provision of water and sewerage services in american cities, 1880-1925. *Journal of economic history*, pages 734–772.

Urban-Brookings Tax Policy Center (2017). State and local finance data: Exploring the census of governments. Technical report.

US Census Bureau (2021). About annual survey of state and local government finances. Technical report.

USDA (2019). High energy cost grants fact sheet. Technical report.

USDA (2020). Solid waste management grants fact sheet. Technical report.

USEPA (1968-1988). Storet legacy.

USEPA (1973). Estimate of municipal wastewater treatment facility requirements. Technical report, Clean Watershed Needs Survey.

USEPA (1980). Handbook of procedures: Construction grants program for municipal wastewater treatment works. Technical report.

USEPA (1997). 1995 community water system survey. Volume 1.

USEPA (1998). How wastewater treatment works. Technical report.

USEPA (2000). Progress in water quality: an evaluation of the national investment in municipal wastewater treatment. Technical report.

USEPA (2001). The national costs to implement tmdls. Technical report.

USEPA (2002). The clean water and drinking water infrastructure gap analysis. Technical report.

USEPA (2004). Primer for municipal wastewater treatment systems. Technical report, Document No. Tech. rep. EPA 832-R-04-001.

USGAO (1994). Water pollution: Information on the use of alternative wastewater treatment systems. Technical report.

Vegh, C. and Vuletin, G. (2015). Unsticking the flypaper effect in an uncertain world. *Journal of Public Economics*, 131.

Watson, T. (2006). Public health investments and the infant mortality gap: Evidence from federal sanitation interventions on us indian reservations. *Journal of Public Economics*, 90(8-9):1537–1560.

Yinger, J., Bloom, H. S., and Boersch-Supan, A. (2016). *Property taxes and house values: The theory and estimation of intrajurisdictional property tax capitalization*. Elsevier.

Zimmer, R., Henry, G. T., and Kho, A. (2017). The effects of school turnaround in tennessee's achievement school district and innovation zones. *Educational Evaluation and Policy Analysis*, 39(4):670–696.