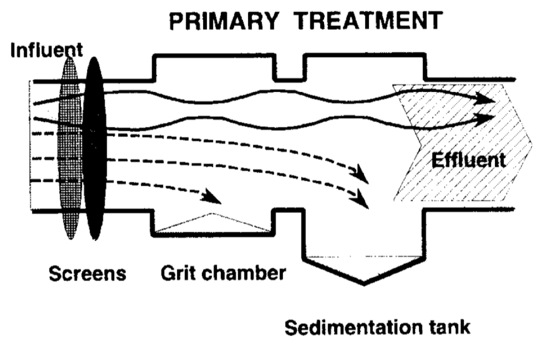
Online Appendix for:

A Watershed Moment: The Clean Water Act and Birth Weight

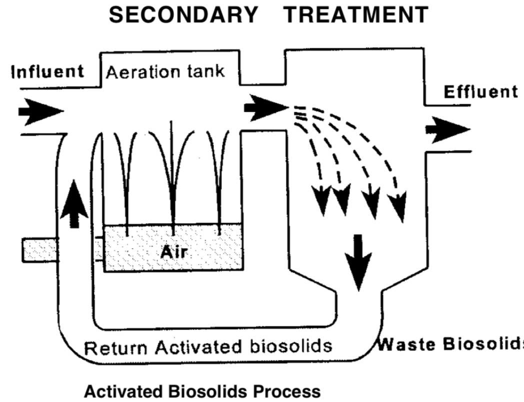
Authors:
Patrick Flynn
Michelle Marcus

A Appendix: Additional Results

A.1 Additional Figures



(a)

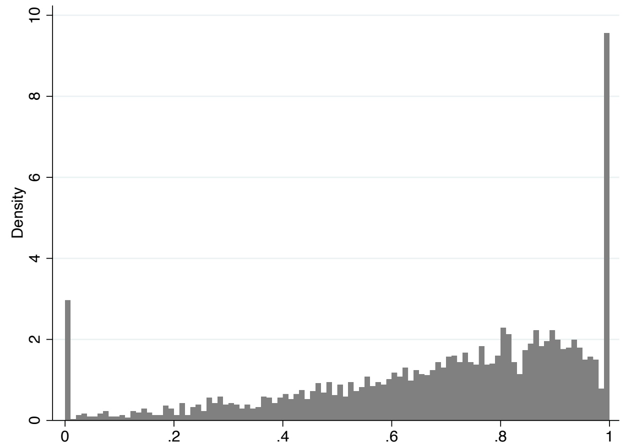


(b)

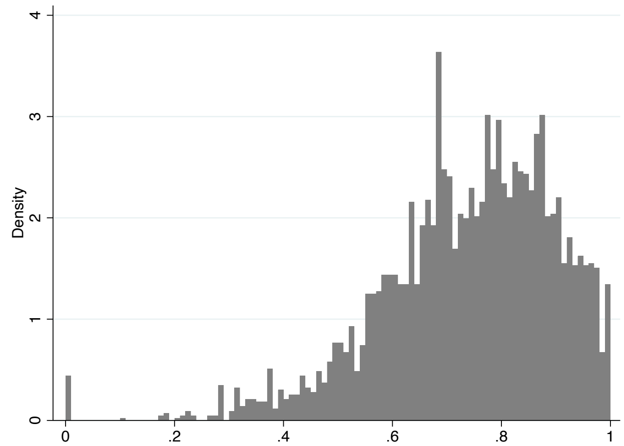
Figure A1: Primary vs Secondary Treatment Technology

Source: USEPA (1998)

Figure A2: Timing of First Grant

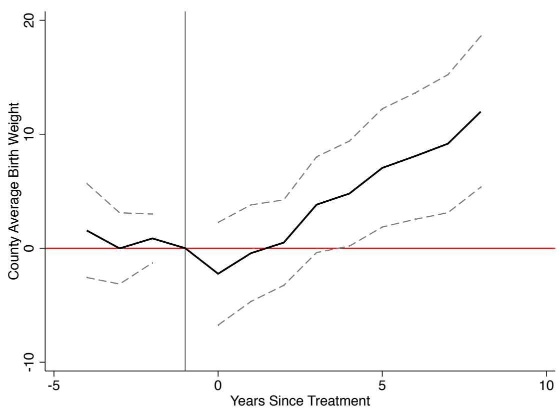


(a) County Level Treatment

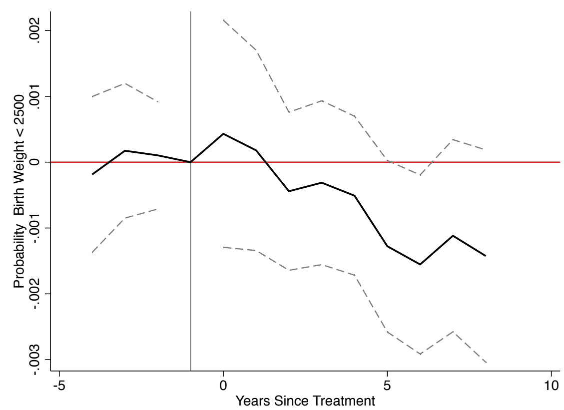


(b) Facility Level Treatment

Figure A3: Percent of Population Living Within a Mile of a Treated Waterway in 1988



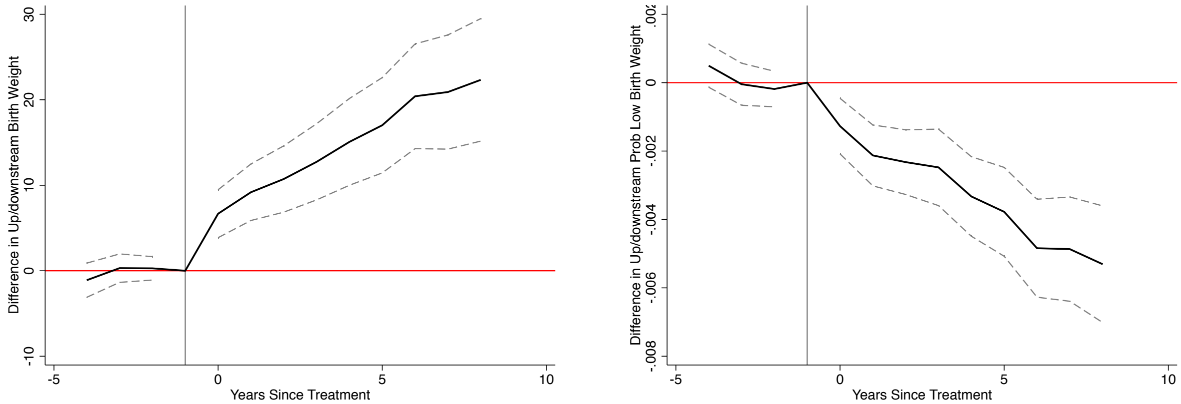
(a) Birth Weight



(b) Probability of Low Birth Weight

Figure A4: Birth Outcomes Downstream from Grant Facilities

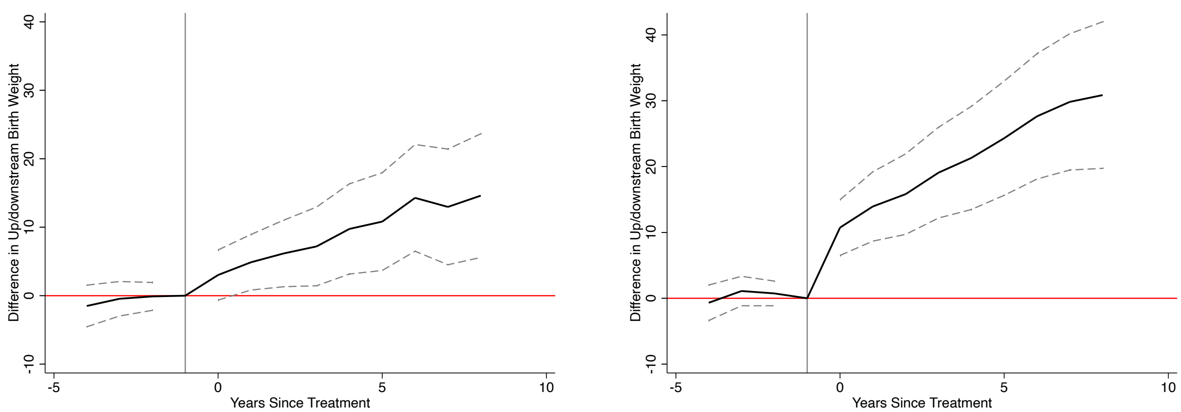
Notes: These figures plot the π_t and γ_t from estimating $Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} * pct_{cy} + \beta X_{cy} + \alpha_c + \alpha_y + \epsilon_{cy}$. pct_{cy} is a continuous variable that takes values from zero to one, and indicates 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 in sub-figure (a), and the probability of being born weighing less than 2500 grams in county c in year y in sub-figure (b).
Source: National Center for Health Statistics (1968-1988a)



(a) Birth Weight (b) Probability of Low Birth Weight

Figure A5: Difference in Birth Outcomes Up and Downstream from Grant Facilities

Notes: These figure plot the π_t and γ_t from estimating $\Delta Y_{py} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_p^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_p^* = t\} * pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py}$. 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 respectively, as well as controls for the percent of up and downstream counties' births of a given birth order, and averages of up and downstream mother's age and race and child gender. The estimates are weighted by total number of births in counties up and downstream from facility p in year y. The dependent variable is the difference in birth weight between up and downstream counties in year y in sub-figure (a), and the difference in the probability of being born weighing less than 2500 grams between up and downstream counties in year y in sub-figure (b).
Source: National Center for Health Statistics (1968-1988a)



(a) Compliant (b) Non-Compliant

Figure A6: Event Studies by Compliance

Notes: The figures plot the event study estimates from Figure 2b separately in panels (a) and (b) for compliant and non-compliant facilities, t_p. The model includes facility and year fixed effects, α_p and α_y respectively, as well as demographic controls. The dependent variable is the difference in birth weight between up and downstream counties in year y. Source: National Center for Health Statistics (1968-1988a)

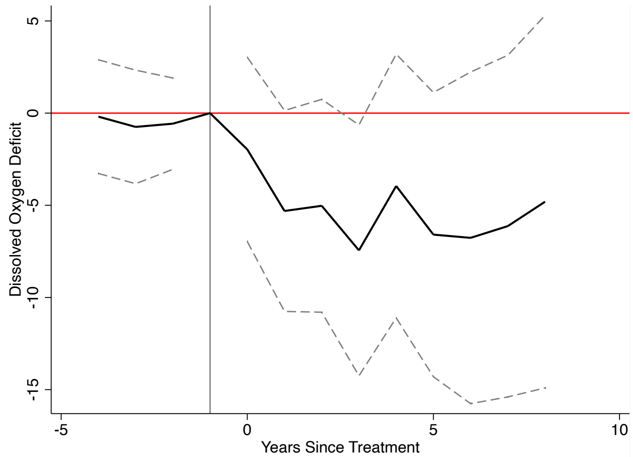


Figure A7: First Stage

A.2 Additional Tables

Table A1: Agents of Waterborne or Water-based Disease

Bacteria	Protozoa	Viruses
Vibrio cholerae	*Giardia lamblia*	Norovirus
Salmonella spp.	*Cryptosporidium parvum*	Sapprovirus
Shigella spp.	*Entamoeba histolitica*	Poliovirus
Toigenic *Escherichia coli*	*Cyclospora cayetanensis*	Coxsackievirus
Campylobacter spp.	*Isospora belli*	Echovirus
Yersinia enterocolitica	Microsporidia	Paraechovirus
Legionella	*Ballantidium coli*	Enteroviruses 69-91
Helicobacter pylori	*Toxoplasma gondii*	Reovirus
	Naegleria fowleri	Adenovirus
		Hepatitis A & E
		Rotavirus
		Astrovirus
		Picobirnavirus
		Coronavirus

Source: Reynolds et al. (2008)

Table A2: 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 bw $<$ 2500	.078	.079	.074	.078	.077	.078	.077
nonwhite	.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 (1968-1988a)

Table A3: Effect on Across the Birth Weight Distribution

	(1) bw $<$ 1000	(2) bw $<$ 1500	(3) 1500 $<$ bw $<$ 2500	(4) bw $>$ 2500
pct pop 1 mile	-0.000320*	-0.000496**	-0.000378	0.000874*
	(0.000174)	(0.000210)	(0.000439)	(0.000523)
Observations	64239	64239	64239	64239

Standard errors in parentheses

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

Notes: This table presents re-estimates our difference-in-differences results from Column 2 of Table 3 on different bins of birth weight. Column 1 shows the effect on Extremely Low Birth Weight (ELBW), defined as births below 1,000 grams. Column 2 shows the effect on Very Low Birth Weight, defined as births below 1,500 grams. Column 3 shows the effect on births between 1,500 and 2,500 grams, which includes births classified as Low Birth Weight but not VLBW or ELBW. Column 4 shows the effect on births above 2,500 grams.

Table A4: First Stage

	(1) non-compliant	(2) compliant	(3) DDD	(4) DDD	(5) DDD
pct pop 1 mile	-3.946***	-0.132	-0.132	-0.271	0.659
	(1.496)	(1.908)	(1.908)	(0.767)	(1.907)
pct pop 1 mile X non-compliant			-3.814	-1.789*	-3.704
			(2.424)	(1.078)	(2.440)
demographic controls	X	X	X	X	
unit fixed effects	X	X	X	X	X
year fixed effects	X	X	X	X	X
collapsed to facility level	X	X	X	X	X
weighted	X	X	X		X
N	12201	11378	23579	23579	23579

Standard errors in parentheses

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

Table A5: Alternative Small Bandwidths

	(1) 25 miles downstream .5 mile buffer	(2) 25 miles downstream 1.5 mile buffer
	county average birth weight	
pct pop .5 miles	10.70**	
	[1.961,19.44]	
pct pop 1.5 miles		6.621**
		[1.081,12.16]
demographic controls	X	X
unit and year fixed effects	X	X
collapsed to facility level	X	X
N	82320	82320

95% confidence intervals in brackets

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

Notes: This table presents (weighted) estimates from the following model: $\Delta Y_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \epsilon_{py}$. 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 A6: Drop Top Quartile of Land Area

	(1)
pct pop 1 mile	5.429**
	(2.515)
demographic controls	X
unit and year fixed effects	X
collapsed to county level	X
observations	48174

Standard errors in parentheses

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

Notes: This table reproduces difference-in-difference estimates from Column 2 of Table 3 after dropping the counties in the top quartile of land area (i.e. counties with the largest geographic area).

Table A7: Correlation of Treatment Variables

	(1)
	pct pop public water
pct downstream	0.927***
	(0.00944)
Observations	8463

Standard errors in parentheses

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

Notes: This table shows the correlation between the percent of the population living in a treated public water system's service area and the percent of the population living within a mile of a treated waterway by presenting estimates from the following model: $pws_{cy} = \beta pct_{cy}$ where pws_{cy} is a variable that takes values between zero and one and indicates the proportion of county population population living in a treated public water system's service area.

A.3 Heterogeneity

We examine the heterogeneity of our estimates across race in Table A8 by estimating equation 4 on sub-samples of white and non-white births from counties with sizable non-white populations.[27] 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

[27]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.

especially true if we think there is a convex relationship between pollution and health.

We address this in columns 3 and 4 of Table A8. In column 3, we drop all observations from facilities that received a grant after 1976 and re-estimate equation 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.

Table A8: Heterogeneous Effects

	(1) white	(2) nonwhite	(3) early grants	(4) later grants
pct downstream X non-compliant	11.37***	14.32	14.04**	11.95**
	[3.778,18.97]	[-7.037,35.68]	[1.241,26.84]	[1.422,22.48]
demographic controls	X	X	X	X
unit and year fixed effects	X	X	X	X
collapsed to facility level	X	X	X	X
N	35406	35406	51639	31080

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

Notes: This table re-estimates the equation 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 (1968-1988a)

A.4 Mortality

Using data from National Center for Health Statistics (1968-1988b), we re-estimate equation 4 with mortality as the dependent variable in Table A9. 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.

Table A9: 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 downstream X non-compliant	0.389	10.11	-14.51	-3.723	-35.34	-19.66	1.607
	[-19.65,20.43]	[-10.01,30.23]	[-63.08,34.06]	[-43.27,35.82]	[-119.9,49.17]	[-68.25,28.93]	[-8.503,11.72]
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
N	82320	82320	82320	82320	82320	82320	82320

95% confidence intervals in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Notes: This table presents (weighted) estimates from the following model: $\Delta Y_{py} = \gamma pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py}$. 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 (1968-1988b); Solley et al. (1988)

A.5 Public Water

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 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.[28]

We show that our results are driven by counties that had some public water systems that drew from surface water sources in Table A10. Column 1 of Table A10 re-estimates equation 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.[29]

We disaggregate these results further in Table A11 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 A12. 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

[28] We use data from 1985 because it is the earliest year for which information on county level water usage is available.

[29] Columns 6 and 7 of Table A2 suggest that communities served by surface and groundwater systems serve similar populations.

maps of public water supply areas from 8 states (see Section C.2 for details on this data). This reduces the estimation sample by 86% (from 64,239 to 8,463 county-year observations). 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 significant at the 10 percent level and statistically indistinguishable from our main estimates.[30] We showed that this treatment measure is correlated with the percent of the population living within a mile of a treated waterway for theses eight states in Table A7, which suggests that some of our main results are driven by this public water channel.

Note that, if populations are receiving publicly provided drinking water from other counties, our county-level measure of treatment may not accurately describe treated populations. We do not have data on the locations of public water system's source wells, but, while water service areas and county borders do not always perfectly align, community water systems generally serve areas no larger than counties (USEPA, 1997).

Table A10: Effects by Public Water Source

	Surface Water	Ground Water
	(1)	(2)
Panel A	county average birth weight	
pct downstream X non-compliant	8.893**	-5.137
	[1.874,15.91]	[-21.34,11.06]
Panel B	probability birth weight < 2500 grams	
pct downstream X non-compliant	-0.000952	0.000132
	[-0.00261,0.000705]	[-0.00375,0.00401]
demographic controls	X	X
unit and year fixed effects	X	X
collapsed to facility level	X	X
N	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 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 (1968-1988a); Solley et al. (1988)

[30] While estimates in Table A12 are slightly smaller than those from the full sample, we re-estimate equation 2 on the sample of states that we have public water supply data for in Table A13 and find similar estimates for this reduced sample with our main specification.

Table A11: Public Water Source Triple Difference

	Surface	Ground	DDD
	(1)	(2)	(3)
Panel A. Non-compliant	county average birth weight		
pct downstream	10.15***	-7.879	-7.879
	[5.927,14.38]	[-20.35,4.597]	[-20.23,4.473]
pct downstream X surface			18.03***
			[4.976,31.09]
N	30009	4200	34209
Panel B. Non-compliant	probability birth weight $<$ 2500 grams		
pct downstream	-0.000872*	0.00103	0.00103
	[-0.00182,0.0000796]	[-0.00192,0.00399]	[-0.00189,0.00396]
pct downstream X surface			-0.00190
			[-0.00498,0.00117]
N	30009	4200	34209
Panel C. Compliant	county average birth weight		
pct downstream	3.111	3.110	3.110
	[-0.861,7.083]	[-4.426,10.65]	[-4.402,10.62]
pct downstream X surface			0.000404
			[-8.497,8.498]
N	37023	11088	48111
Panel D. Compliant	probability birth weight $<$ 2500 grams		
pct downstream	-0.000333	-0.00183	-0.00183
	[-0.00138,0.000714]	[-0.00419,0.000522]	[-0.00418,0.000515]
pct downstream X surface			0.00150
			[-0.00107,0.00407]
N	37023	11088	48111
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to facility level	X	X	X

95% confidence intervals in brackets

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

Notes: This table describes the effects of Clean Water Act grants on birth outcomes depending on public water source. Column 1 estimates $\Delta Y_{py} = \gamma pct_{py} + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py}$ 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: $\Delta Y_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * s_p + \beta X_{py} + \phi X_{py} * s_p + \alpha_y * s_p p + \alpha_p + \alpha_y + \epsilon_{py}$ 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 (1968-1988a); Solley et al. (1988)

Table A12: Exposure Defined by Percent on Public Water Supply

	(1) birth weight	(2) prob bw < 2500
pct pop public water	4.705*	-0.000224
	[-0.411,9.821]	[-0.00210,0.00165]
demographic controls	X	X
unit and year fixed effects	X	X
collapsed to county level	X	X
N	8463	8463

95% confidence intervals in brackets

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

Notes: In this table, we re-estimate the results in column 2 of Table 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 (1968-1988a)

Table A13: Limit Sample to States with Public Water Supply Maps

	(1) birth weight	(2) prob bw < 2500
pct downstream	2.242	-0.000626
	[-4.038,8.522]	[-0.00268,0.00143]
demographic controls	X	X
unit and year fixed effects	X	X
collapsed to county level	X	X
N	8463	8463

95% confidence intervals in brackets

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

Notes: In this table, we re-estimate the results in column 2 of Table 3 on the eight states that we have public water supply data for.
Source: National Center for Health Statistics (1968-1988a)

A.6 Quantifying the Benefits to Infant Health

Our estimates suggest that 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. We use this information to quantify the benefits to one measure of health. We note that a full accounting of the health benefits would include many other infant health measures that are often studied in the literature (e.g., gestation length, SGA, neonatal mortality, maternal complications, etc.), as well as child, adolescent, and adult health measures.

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 3). Given the measurement error that may be incorporated into our estimates due to data constraints, we think it is most helpful to use the upper bound of the confidence interval we estimate for low birth weight.

In terms of the costs associated with low birth weight, 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 Bharadwaj et al. (2018) finds that low birth weight reduces permanent labor income by 3.4 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 rough back-of-the-envelope estimate.

While our estimates face measurement error that may attenuate the effects and a more comprehensive calculation of the health benefits of the CWA would include other potentially impacted health outcomes, we estimate that the upper-bound of the confidence interval on the effects on low birth weight generates benefits of about \$32 billion. This is about 21 percent of the amount needed to make the CWA cost effective.[31] Future work should consider the effect of the CWA on additional measures of health to provide a more comprehensive cost-benefit analysis.

B Appendix: Robustness

B.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 4 defining treated waterways as those either 5 or 10 miles downstream from a treated facility in Table B1. The results are similar to those presented in Section 4.

[31] In total, CWA grants to wastewater treatment facilities cost an estimated \$153 billion (in 2014 dollars).

Table B1: Other Distances Downstream

	non-compliant	compliant	DDD
	(1)	(2)	(3)
Panel A. 5 miles downstream	county average birth weight		
pct downstream	14.68***	6.358***	6.358***
	[9.192,20.18]	[2.190,10.53]	[2.191,10.52]
pct downstream X non-compliant			8.326**
			[1.435,15.22]
N	35973	50379	86352
Panel B. 10 miles downstream	county average birth weight		
pct downstream	14.44***	6.167***	6.167***
	[8.986,19.90]	[2.023,10.31]	[2.024,10.31]
pct downstream X non-compliant			8.278**
			[1.429,15.13]
N	35154	49413	84567
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to facility level	X	X	X

95% confidence intervals in brackets

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

Notes: This table presents (weighted) estimates from the following model: $bw_{py} = \gamma_0^{DD} pct_{py} + \gamma^{DDD} pct_{py} * t_p + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \epsilon_{py}$. 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 (1968-1988a)

B.2 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, 2021). 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. Note that 93 counties out of 3,064 total counties were never-treated. We can then estimate the following stacked difference-in-difference:

$$Y_{py} = \gamma^{stacked} pct_{py} + \alpha_{ps} + \alpha_{sy} + \epsilon_{psy} \tag{5}$$

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 5 in Table B2. 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.

Table B2: Stacked Difference in Difference

	(1) not yet treated	(2) never treated	(3) both
Panel A	county average birth weight		
pct downstream	5.209**	26.96***	5.458**
	[0.247,10.17]	[19.12,34.80]	[0.509,10.41]
Panel B	probability bw < 2500		
pct downstream	-0.00134**	-0.00541***	-0.00139**
	[-0.00243,-0.000255]	[-0.00705,-0.00377]	[-0.00247,-0.000308]
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to facility level	X	X	X
N	83580	63041	86088

95% confidence intervals in brackets

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

Notes: This table presents results from estimating the following stacked difference in difference: $Y_{py} = \gamma^{stacked} pct_{psy} + \alpha_{ps} + \alpha_{sy} + \epsilon_{psy}$. 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 (1968-1988a)

B.3 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 (Callaway et al., 2021), 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.

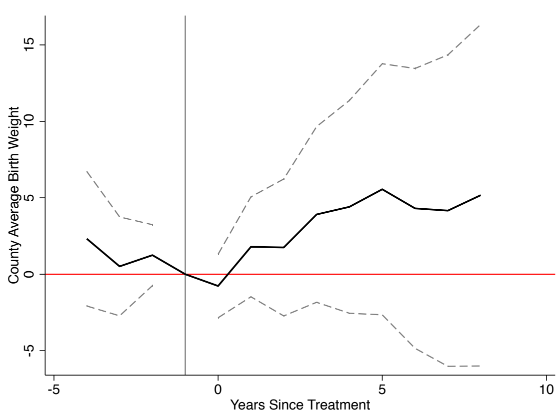
We first estimate the following event study

$$Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} + \beta X_{cy} + \alpha_c + \alpha_y + \epsilon_{cy} \tag{6}$$

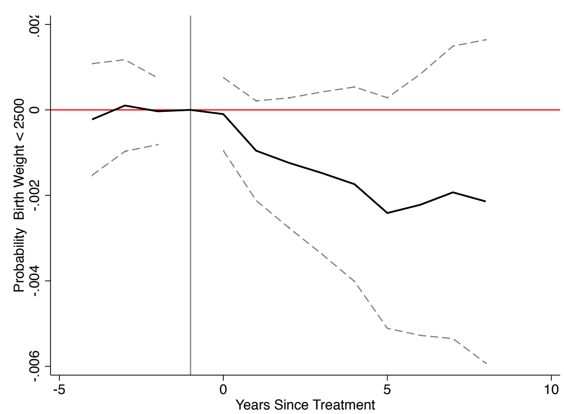
We present estimates of equation 6 with average birth weight and the probability of low birth weight in Figure B1. The shapes of these event studies are similar to those in the main text.

When we define treatment with a dummy variable, we can deal with the problems caused by dynamic treatment effects discussed in Section B.2 in a more sophisticated way. To summarize these event studies, we use Callaway and Sant'Anna (2020) to estimate treatment effects in Table B3.

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.



(a) Birth Weight



(b) Probability of Low Birth Weight

Figure B1: Birth Outcomes Downstream from Grant Facilities (Binary Treatment)

Notes: These figures plot the π_t and γ_t from estimating $Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} + \beta X_{cy} + \alpha_c + \alpha_y + \epsilon_{cy}$. 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 in sub-figure (a), and the probability of being born weighing less than 2500 grams in county c in year y in sub-figure (b).
Source: National Center for Health Statistics (1968-1988a)

Table B3: Callaway and Sant'Anna (2020) Estimates

	birth weight	prob bw < 2500
	(1)	(2)
grant X downstream	4.85*	-0.0018
	(2.60)	(0.0032)
N	64239	64239

standard errors in parenthesis

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

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 (1968-1988a)

B.4 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 B4 by estimating equation 7.

$$\Delta Y_{py} = \gamma pct_{py} + \eta pct_{py} * t_p + \pi pct_{py} * Interact_p + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py} \tag{7}$$

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.[32] All other variables are defined analogously to those in equation 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.

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

Table B4: Other Interactions

	(1)	(2)	(3)	(4)
	county average birth weight			
pct downstream X non-compliant	6.464**	5.268**	5.389	6.736
	[0.664,12.26]	[0.143,10.39]	[-2.149,12.93]	[-2.078,15.55]
pct downstream	4.719*	7.304***	5.888	5.687
	[-0.507,9.945]	[2.763,11.84]	[-1.797,13.57]	[-2.950,14.32]
pct downstream X total flow	-0.0263			0.0347
	[-0.0652,0.0126]			[-0.0314,0.101]
pct downstream X population served		-0.00000700		-0.0000165
		[-0.0000184,0.00000441]		[-0.0000377,0.00000467]
pct downstream X other modification			-0.903	-2.871
			[-14.13,12.33]	[-16.76,11.02]
demographic controls	X	X	X	X
unit and year fixed effects	X	X	X	X
collapsed to facility level	X	X	X	X
N	35049	45864	30597	24717

95% confidence intervals in brackets

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

Notes: This table estimates $\Delta Y_{py} = \gamma pct_{py} + \eta pct_{py} * t_p + \pi pct_{py} * Interact_p + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py}$ 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 3.
Source: National Center for Health Statistics (1968-1988a)

B.5 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 2b without binning these unbalanced endpoints in Figure B2. 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 after treatment, consistent with projects taking up to 10 years from grant application to project completion (USEPA, 2002).

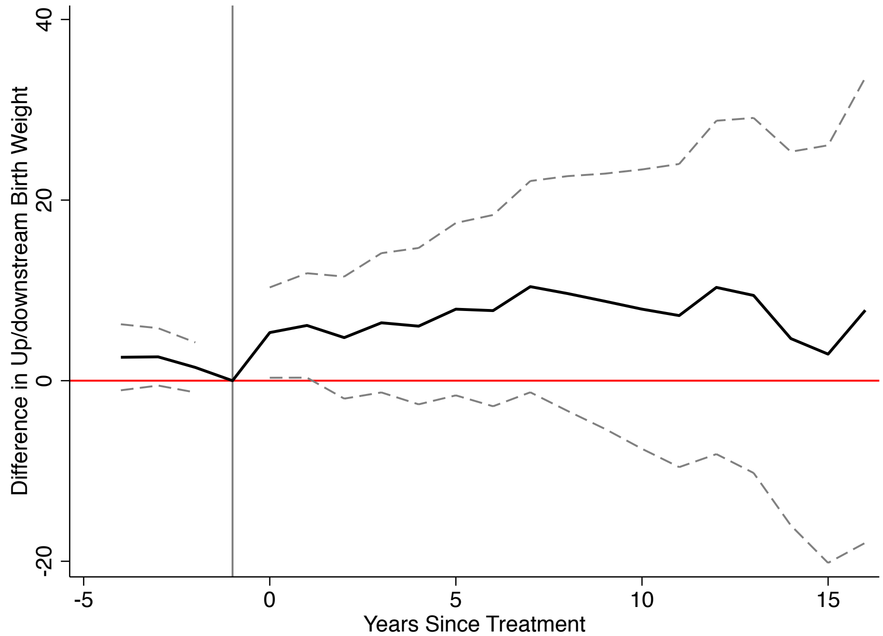


Figure B2: Birth Weight Triple Difference

Notes: These figures plot the θ_t and η_t from estimating $\Delta Y_{py} = \sum_{t=-5}^{-2} \theta_t 1\{y-y_p^* = t\} * t_p + \sum_{t=0}^{16} \eta_t 1\{y-y_p^* = t\} * pct_{py} * t_p + \sum_{t=-4}^{-2} \pi_t 1\{y-y_p^* = t\} + \sum_{t=0}^{16} \gamma_t 1\{y-y_p^* = t\} * pct_{py} + \beta X_{py} + \phi X_{py} * t_p + \alpha_y * t_p + \alpha_p + \alpha_y + \epsilon_{py}$. All variables are defined analogously to those in Figure 2. The dependent variable is the difference in birth weight between up and downstream counties in year y.
Source: National Center for Health Statistics (1968-1988a)

C Appendix: Additional Data Details

C.1 Birth Data Details

C.1.1 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 C1.

Table C1: 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

C.1.2 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 C2 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 2b on this sample in Figure C1. We then re-estimate the results presented in Table 3 on this sample and report the results in Table C3, which are similar to those reported in Section 4.

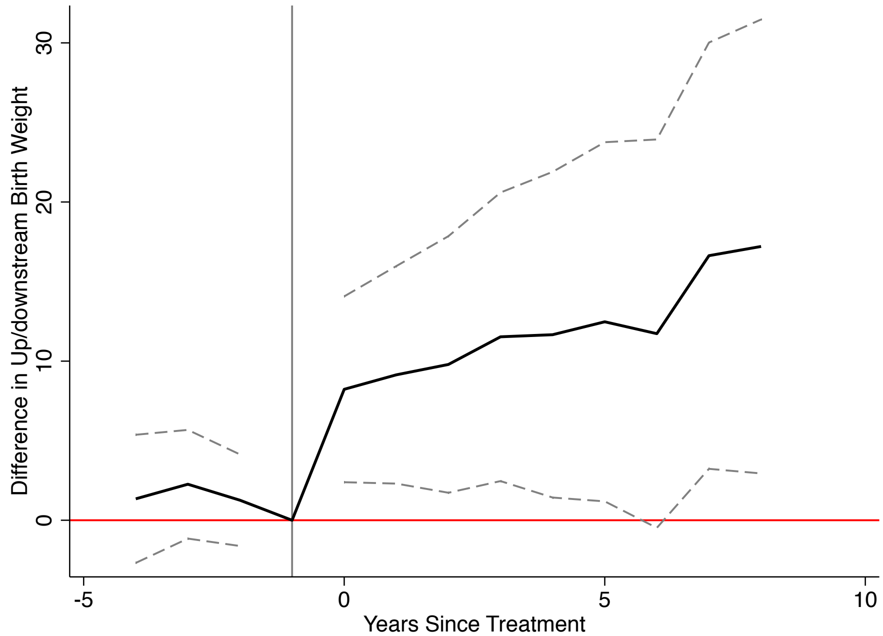


Figure C1: Birth Weight Triple Difference: Random Sample

Notes: This Figure re-estimates the results in Figure 2b 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 C2.
Soure: National Center for Health Statistics (1968-1988a)

Table C2: 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 C3: Triple Difference: Random Sample

	(1) non-compliant	(2) compliant	(3) DDD
pct downstream	12.38***	4.448**	4.448**
	[7.015,17.74]	[0.303,8.593]	[0.304,8.592]
pct downstream X non-compliant			7.933**
			[1.157,14.71]
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to facility level	X	X	X
N	34188	48132	82320

95% confidence intervals in brackets

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

Notes: This table re-estimates the specifications in Columns 5-7 in Panel A of Table 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 (1968-1988a)

C.2 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 (CT 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 (NJDEP, 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 (NCDEQ, 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 (PASDA, 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).

C.3 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.[33] Question 21c then asks whether a facility is currently in compliance with both the EPAs secondary treatment mandate and any higher mandates.[34] 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.

[33] In particular, it asks "What level of secondary treatment must the discharge of this pants meet by July 1, 1977? 1. Secondary treatment level as defined by the EPA, OR 2. Higher level of secondary treatment required by State."

[34] Question 21c asks "Does the discharge from this plant NOW meet the level of secondary treatment defined in 21b? 1. Yes, 2. No."