A Appendix

A.1 Recreation Results

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.[25]

First, we test this channel in Table A1 by separately estimating equation 3 on subsamples 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 A2, 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., 2018).

[25] 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.

Table A1: Split by Recreational Spending Per Capita

	(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***
	[-7.608,7.934]	[-1.908,11.31]	[6.503,23.88]
Panel B	probability birth weight < 2500 grams		
pct pop 1 mile	-0.000429	-0.000453	-0.00220***
	[-0.00193,0.00107]	[-0.00234,0.00143]	[-0.00385,-0.000561]
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to county level	X	X	X
N	21147	20160	22617

95% confidence intervals in brackets

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

Notes: This table re-estimates the specification in column 2 of Table 4 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 (1968-1988a); Bureau of Economic Analysis (2012-2017)

Table A2: Triple Difference Split by Recreational Spending Per Capita

	(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
	[-9.664,12.13]	[3.758,25.76]	[-7.433,15.46]
Panel B	probability birth weight < 2500 grams		
pct pop 1 mile X non-compliant	0.00183	-0.00352**	0.000262
	[-0.000580,0.00423]	[-0.00687,-0.000171]	[-0.00251,0.00303]
demographic controls	X	X	X
unit and year fixed effects	X	X	X
collapsed to facility level	X	X	X
N	20748	19656	23247

95% confidence intervals in brackets

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

Notes: This table re-estimates the specification in column 3 of Table 5 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 (1968-1988a); Bureau of Economic Analysis (2012-2017)

A.2 Mortality

Using data from National Center for Health Statistics (1968-1988b), we re-estimate equation 5 with mortality as the dependent variable in Table A3. 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 A3: 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
	[-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.3 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 5 defining treated waterways as those either 5 or 10 miles downstream from a treated facility in Table A4. The results are similar to those presented in Section 4.

Table A4: 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***
	[9.192,20.18]	[2.190,10.53]	[2.191,10.52]
pct pop 1 mile X non-compliant			8.326**
			[1.435,15.22]
N	35973	50379	86352
Panel B. 10 miles downstream	county average birth weight		
pct pop 1 mile	14.44***	6.167***	6.167***
	[8.986,19.90]	[2.023,10.31]	[2.024,10.31]
pct pop 1 mile 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)

A.4 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 A5 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 5 with different bandwidths. In column 1, this bandwidth is .5 miles, and in column 2, it is 1.5 miles.

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.

A.5 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, 2019). 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:

$$Y_{py} = \gamma^{stacked} pct_{py} + \alpha_{ps} + \alpha_{sy} + \epsilon_{psy} \tag{7}$$

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 7 in Table A6. 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 A6: Stacked Difference in Difference

	(1)	(2)	(3)
	not yet treated	never treated	both
Panel A	county average birth weight		
pct pop 1 mile	5.209**	26.96***	5.458**
	[0.247,10.17]	[19.12,34.80]	[0.509,10.41]
Panel B	probability bw < 2500		
pct pop 1 mile	-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)

A.6 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

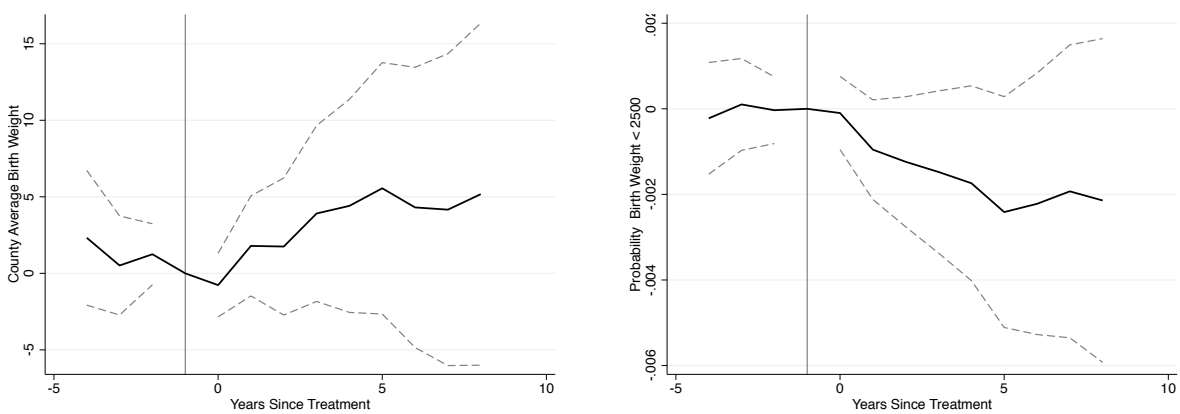
$$Y_{cy} = \sum_{t=-5}^{-2} \pi_t 1\{y - y_c^* = t\} + \sum_{t=0}^{9} \gamma_t 1\{y - y_c^* = t\} + \beta X_{cy} + \alpha_c + \alpha_y + \epsilon_{cy} \tag{8}$$

which is identical to equation 2 except the timing dummies are not interacted with the percent of county population living within a mile of a treated waterway.

We present estimates of equation 8 with average birth weight and the probability of low birth weight in Figure A1. 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 A.5 in a more sophisticated way. To summarize these event studies, we use Callaway and Sant'Anna (2020) to estimate treatment effects in Table A7.

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 A1: 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 A7: 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)

A.7 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 A8 by estimating equation 9.

$$\Delta Y_{py} = \gamma pct_{py} + \eta pct_{py} * t_p + \pi pct_{py} * Interact_p + \beta X_{py} + \alpha_p + \alpha_y + \epsilon_{py} \tag{9}$$

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.[26] All other variables are defined analogously to those in equation 4.

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.

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

Table A8: 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
	[0.664,12.26]	[0.143,10.39]	[-2.149,12.93]	[-2.078,15.55]
pct pop 1 mile	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 pop 1 mile X total flow	-0.0263			0.0347
	[-0.0652,0.0126]			[-0.0314,0.101]
pct pop 1 mile X population served		-0.00000700		-0.0000165
		[-0.0000184,0.00000441]		[-0.0000377,0.00000467]
pct pop 1 mile 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 4.
Source: National Center for Health Statistics (1968-1988a)

A.8 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 6a without binning these unbalanced endpoints in Figure A2. 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).

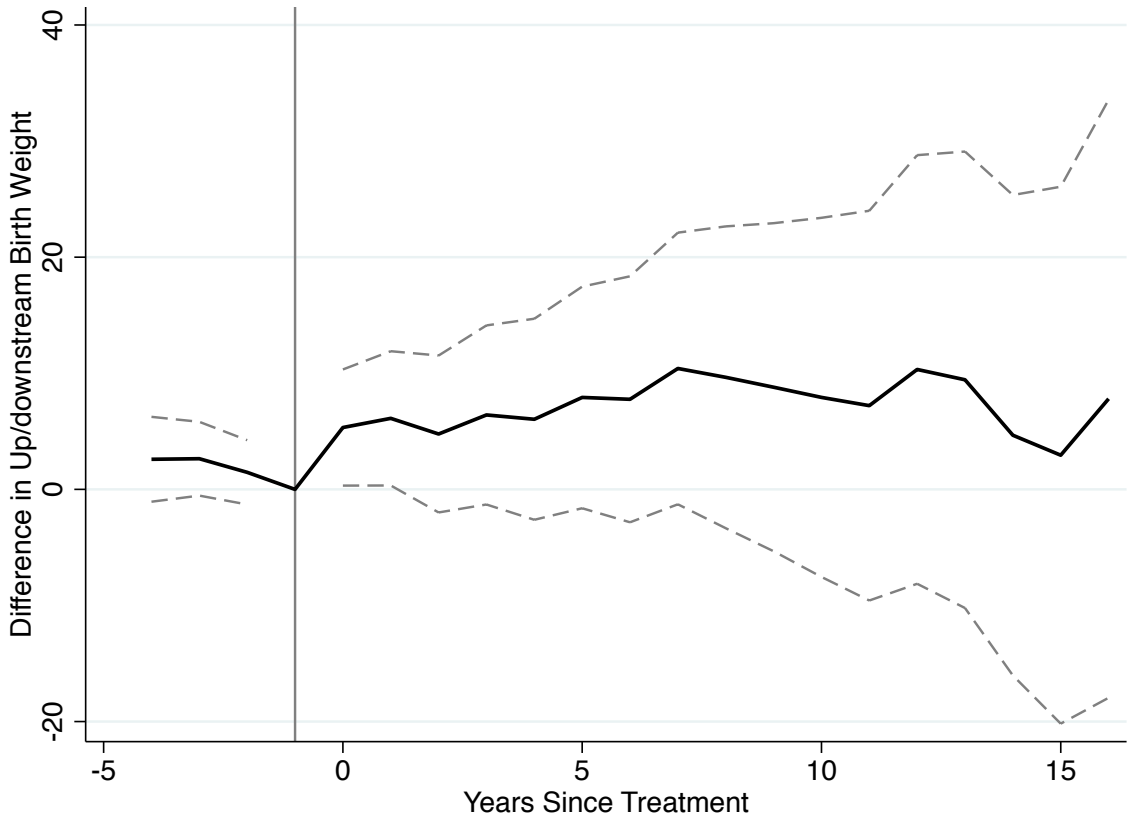


Figure A2: Birth Weight Triple Difference

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 6. 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)

A.9 Birth Weight Data Details

A.9.1 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 A9.

Table A9: 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

A.9.2 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 A10 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 equation 6 on this sample in Figure A3. We then re-estimate the results presented in Table 5 on this sample and report the results in Table A11, which are similar to those reported in Section 4.

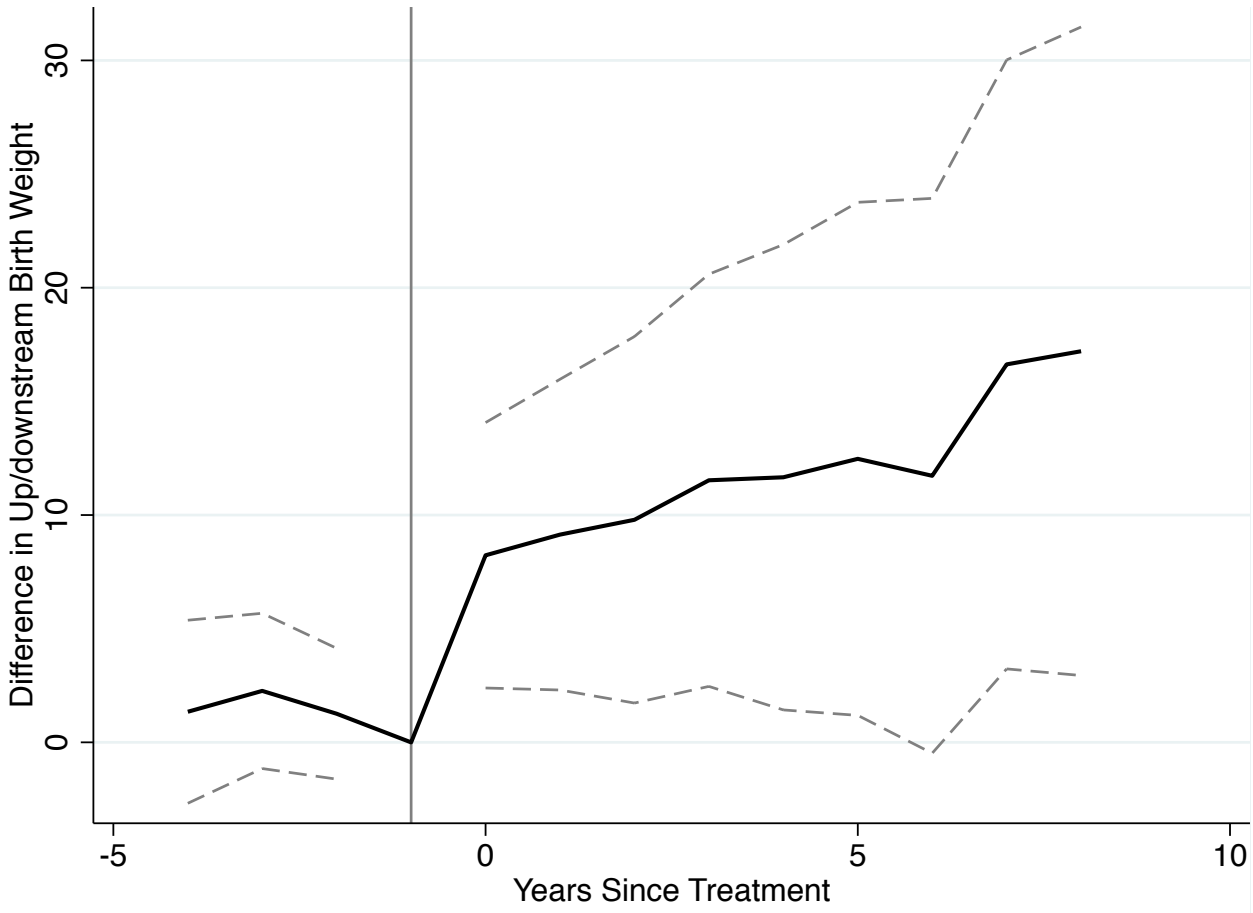


Figure A3: Birth Weight Triple Difference: Random Sample

Notes: This Figure re-estimates the results in Figure 6a 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 A10.
Soure: National Center for Health Statistics (1968-1988a)

Table A10: 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 A11: Triple Difference: Random Sample

	(1) non-compliant	(2) compliant	(3) DDD
pct pop 1 mile	12.38***	4.448**	4.448**
	[7.015,17.74]	[0.303,8.593]	[0.304,8.592]
pct pop 1 mile 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 Panel A of Table 5 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)

A.10 Public Water Supply

In the main text, we argue that the effect of CWA grants on birth weight is driven by reduced contamination of publicly provided water. We provide further evidence of this in Table A12. In this table, we re-estimate equation 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. We calculate this using maps of public water supply areas from 8 states (see Section A.10.1 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.

Table A12: 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 4 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)

A.10.1 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).