Private provision of public goods by environmental groups

Laura Grant[a,1] and Christian Langpap[b]

[a]Robert Day School of Economics and Finance, Claremont McKenna College, Claremont, CA 91711; and [b]Department of Applied Economics, Oregon State University, Corvallis, OR 97331

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved August 17, 2018 (received for review May 17, 2018)

Many environmental nonprofit groups are assumed to provide public goods. While an extensive literature examines why donors join and give to nonprofits, none directly tests whether donations actually provide public goods. We seek such a test by using a common form of environmental organization: watershed groups. We find their increased presence resulted in lower dissolved oxygen deficiency and higher proportions of swimmable and fishable water bodies. Increased donations to and expenditures by the groups also improved water quality. Thus, private groups likely played a role in mitigating environmental problems. Overall, our results indicate private provision of a public good by nonprofit organizations.

environmental groups | nonprofits | private provision | public goods | water quality

Public goods are characterized by inadequate provision; there are numerous examples of shortfalls. For instance, the United Nations Sustainable Development Goals reflect unmet targets for clean water and sanitation, biodiversity, and climate change mitigation. Shortfalls occur because individuals often choose immediate personal benefit, not collective long-term gain. Furthermore, individuals have an incentive to free ride because each person benefits from everyone else's contribution.

Collective action through group formation may overcome the inertia of underprovision of public goods (1). Nonprofit groups may be a mechanism for individuals to improve coordination and decision making toward mitigating collective action problems. Over one million nonprofits operate in the United States, many of which aim to provide public goods. In 2016, these nonprofits received almost $282 billion in donations from individuals (2).

The question of whether nonprofit groups increase the level of provision of public goods or the efficiency with which they are provided is of long-standing interest (e.g., refs. 3 and 4). However, despite the large volume of donations and number of organizations, there are no direct empirical tests of whether these organizations provide public goods. The literature focuses on explaining why individuals donate to nonprofit organizations. Theoretical work establishes donor motivations of altruism, warm glow, social approval, public prestige, or a desire to signal income (5–10). Empirical studies examine how household characteristics, fundraising expenditures, tax incentives, and social information influence donations, as well as whether private funding substitutes or complements public funding (11–19).

This literature does not directly measure outputs, the supply of the public good. Instead, it models the amount of the public good simply as the sum of individual inputs, the donations to nonprofit organizations (9). Examples include public radio (11), rural health care (13), green electricity programs (16), or charitable giving broadly defined (12, 14, 15, 20).

Measuring inputs rather than outputs misses how nonprofits generate the public good and how much of it they produce. If donors care about providing public goods, donations are only meaningful when they increase the supply of the public good (21). However, there is a growing literature that analyzes the structure and behavior of nonprofits and identifies several characteristics of these organizations that may reduce their effectiveness at converting donations into public goods.

Nonprofits lack the traditional incentives of for-profit firms because no one has claim over residual earnings. Nonprofits also depend largely on noncontractible donations and have managers whose objectives may not match the goals of donors or society as a whole. Furthermore, the beneficiaries of public goods supplied by nonprofits cannot rely on market forces to penalize or reward these organizations. These factors can incentivize managers to spend residual earnings on perquisite consumption, while reducing program operations (22–25). Additionally, competition for donors increases fundraising and nonprogram expenses, thereby potentially reducing production of the public good (24, 26, 27).

These stylized facts about nonprofits raise the possibility that more donations to these organizations may not translate proportionally into more provision of public goods. Hence, our primary contribution is to empirically examine the relationship between nonprofits and public good provision. A few articles evaluate the efficacy of international development nongovernmental organizations providing foreign aid (28, 29).* However, these articles use indirect measures of public good provision without relating inputs to outputs.

This article analyzes the relationship between nonprofits and public goods by examining environmental nonprofit groups that focus on providing and protecting environmental amenities (30–33). There is evidence that environmental groups affect enforcement by environmental regulators (34, 35). However, little is known about whether these groups increase the supply of public goods. We study watershed groups, a common form of environmental organization. These groups raise funds and coordinate the activity of individuals in local communities to protect and restore rivers and other water bodies. We seek evidence that watershed groups are linked to water quality in the United States. Our results indicate a causal relationship between number of water groups and their donations and expenditures, and improvements in water quality during the study period. We

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Economics, Environment, and Sustainable Development," held January 17–18, 2018, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/economics-environment-and.

Author contributions: L.G. and C.L. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: lgrant@cmc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1805336115/-/DCSupplemental.

Published online October 3, 2018.

*Wane W (2004) The quality of foreign aid: Country selectivity or donors incentives? (World Bank, Washington, DC). Available at hdl.handle.net/10986/14001. Accessed April 20, 2016.

establish this relationship between inputs and output, yet different improvements in water quality at different locations could lead to large variation in the values of services. Thus, we briefly discuss extrapolations of benefit estimates from the literature, noting that specific valuation of water quality change is beyond the scope of this article.

This analysis combines data on water quality and watershed groups for 2,150 watersheds in the United States over the period 1996–2008. The number of watershed groups across the United States tripled during this period, from 500 to 1,500. Basic inspection shows positive correlation between watershed group activity and impaired waters, suggesting that groups tended to locate where water quality was poor. This selection problem can bias the measurement of impacts of water groups on water quality. Hence, the objective of our empirical strategy is to robustly recover parameters that inform a causal effect of water group activity on water quality. Rather than being concerned with overall model fit, we are interested in the robustness of results across different model specifications. To address this potential bias, we use a matching procedure to create an estimation sample that is similar across observable watershed characteristics, including water quality, at the beginning of our study period. Then we carry out fixed-effects statistical regression on the matched sample.

This article contributes to various strands of literature. We advance the literature on private provision of public goods by focusing on an output resulting from donations to nonprofits. This article also strengthens the growing literature on private enforcement of environmental regulations, which heretofore has focused on litigation. Finally, we contribute to the literature on effectiveness of nonprofit organizations by examining the impact of these groups on water quality.

Results

We use water quality, measured as mean dissolved oxygen deficiency (DOD) in rivers and streams in a watershed for a given year, to quantify public good provision. DOD measures the deficiency in the amount of oxygen dissolved in the water. DOD is the most common and overarching measure of water quality because dissolved oxygen (DO) is critical for many forms of aquatic life that use oxygen in respiration, including fish, invertebrates, bacteria, and plants. DOD is also the water quality measure that has the most data available during our study period.[†,‡] DOD is partly caused by excessive amounts of other common pollutants, such as nitrogen or phosphorus, in the water, which increase nutrient loadings and accelerate eutrophication.

We use three measures of group activity in a watershed in a given year: total number of active groups, total donations to all groups in the watershed, and total expenditures (net of fundraising) by groups in the watershed. Presence and activity of watershed groups can impact water quality in various ways, including oversight and monitoring, direct actions such as organizing volunteers for cleanups or restoration, and indirect actions like advocacy and education.

We use linear regression on a balanced sample to estimate the effect of watershed group activity on DOD. We condition on additional factors that impact water quality at the watershed level: violations of the Clean Water Act (CWA), spending via federal water quality programs, land use, precipitation, election outcomes, population density, per capita income, educational attainment, ethnicity, home ownership, and unemployment. Finally, we include year fixed effects and fixed effects for the state responsible for impairment listing of water bodies in the watershed. We discuss the data and the regression models in more detail in *Methods*.

Our estimates indicate that watershed group activity enhanced the water quality in a watershed during our study period. Table 1 presents estimation results. Coefficients for number of groups, donations, and expenditures are negative and significant. The estimated coefficient for number of water groups suggests that an additional group in a watershed in the preceding year is associated with a DOD reduction of 0.0027 percentage points relative to a watershed with no water groups. This change represents a 1.76% reduction in DOD for the average watershed in a given year. The coefficient for donations indicates that a $100,000 increase in total donations to groups in a watershed (roughly a 10% increase relative to the mean) is associated with a DOD reduction of 0.0043 percentage points. This change represents a 2.81% reduction in DOD for the average watershed. The coefficient for expenditures suggests that a $100,000 increase in total expenditures in a watershed (roughly a 7% increase relative to the mean) is associated with a DOD reduction of 0.0018 percentage points. This change represents a 1.18% reduction in DOD for the average watershed. To provide context for the magnitude of these effects, we compare them with the underlying rate of change in water quality. Over our study period, DOD decreased by 2.6% per year on average (*SI Appendix*, Table S1). Relative to this rate of change, reductions in DOD associated with group activity, between 1.18% and 2.81% in an average year, are meaningful.

The estimated coefficients for other variables follow a priori expectations: CWA violations had a negative effect on water quality. Watersheds with higher per capita income had better water quality, perhaps because areas with higher income tend to be concentrated in more urban watersheds, and watersheds with higher home ownership rates had better water quality, possibly because of impacts of home ownership on social capital, investment in local amenities, coordination, or local political processes that affect water quality. Watersheds with higher percentage of white population had better water quality. Finally, federal water quality programs had a positive impact on water quality. The effects of public and private expenditures are not directly comparable because the federal programs are measured at the state level and the watershed group expenditures are within state, at the watershed level. A rigorous comparison is outside the scope of this article, but a coarse analysis can be carried out by aggregating watershed group expenditures at the state level. This procedure, explained in more detail in *SI Appendix*, suggests that additional federal expenditures had a bigger impact on DOD than additional watershed group expenditures. Given the coarseness of the comparison, this conclusion should be interpreted with caution.

Water groups had a positive effect on water quality, yet the change in average DOD does not have an intuitive interpretation. To gain further insight into the magnitude of these effects, we consider two additional outcomes based on goals from the CWA, which aims to provide for the protection and propagation of fish ("fishable") and for recreation in and on the water ("swimmable"). At present, many bodies of water still do not meet these designated uses.

We expect water group activity to increase the proportion of swimmable and fishable water bodies in a watershed. Estimation results are shown in Table 2. Each additional group increased the proportion of swimmable and fishable by 0.52 and 0.28 percentage points, respectively. For swimmable, this change represents a 0.81% increase relative to the mean. For fishable, it corresponds to a 0.32% increase. An additional $100,000 in donations increased the proportion of swimmable by 0.45 percentage points and the proportion of fishable by 0.34 percentage points. These changes represent a 0.71% and a 0.39% increase relative to the mean for proportion swimmable and fishable, respectively. Expenditures also affect the proportions of swimmable and fishable.

[†]Smith VK, Wolloh CV (2012) Has surface water quality improved since the Clean Water Act? (NBER, Cambridge, MA). Available at www.nber.org/papers/w18192. Accessed November 28, 2017.

[‡]Keiser DA, Shapiro JS (2017) Consequences of the Clean Water Act and the demand for water quality (Department of Economics, Iowa State University, Ames, IA). Available at www.nber.org/papers/w23070. Accessed May 6, 2017.

Table 1. Effects of water groups on water quality: Dissolved oxygen deficiency

Explanatory variables	Effect of groups	Effect of donations	Effect of expenditures
No. of water groups$_{t-1}$	−0.269*** (0.096)		
Donations$_{t-1}$, 1,000s $		−4.3E-05*** (9.5E-06)	
Program expenditures$_{t-1}$, 1,000s $			−1.8E-05*** (3.5E-06)
Violations$_{t-1}$	5.307* (2.848)	5.507* (2.860)	5.431* (2.849)
Federal water quality Expenditures, 1,000s $	-6.8×10^{-05}* (3.6×10^{-05})	-6.9×10^{-05}* (3.6×10^{-05})	-6.9×10^{-05}* (3.6×10^{-05})
Fraction agricultural land	0.056 (3.468)	0.016 (3.445)	0.010 (3.444)
Fraction urban land	6.891 (7.921)	4.842 (7.924)	4.982 (7.974)
Population density, persons/mi^2	0.004 (0.005)	0.005 (0.005)	0.005 (0.005)
Per capita income, 1,000s $	0.001** (0.001)	0.001** (0.001)	0.001** (0.001)
High school education	11.672 (8.967)	12.169 (8.924)	12.320 (8.931)
Unemployment rate	−40.894 (34.684)	−38.492 (34.652)	−38.298 (34.682)
Precipitation, 1,000s mm	0.064 (0.052)	0.058 (0.052)	0.058 (0.052)
Precipitation^2, 1,000s mm^2	-1.8×10^{-04} (2.1×10^{-04})	-1.5×10^{-04} (2.1×10^{-04})	-1.6×10^{-04} (2.1×10^{-04})
Percent Republican vote	−2.968 (3.333)	−2.351 (3.330)	−2.333 (3.331)
Home ownership rate	−21.478*** (6.936)	−21.139*** (6.944)	−21.061*** (6.949)
Percent white population	−16.004*** (5.731)	−15.880*** (5.770)	−15.982*** (5.766)
R^2	0.24	0.24	0.24
Observations	7,204	7,204	7,204

Includes year and state fixed effects. *$P < 0.1$; **$P < 0.05$; ***$P < 0.01$. HUC8-clustered SEs are in parentheses.

An additional $100,000 in expenditures increased the proportion of swimmable by 0.25 percentage points and the proportion of fishable by 0.19 percentage points. These changes represent a 0.38% and a 0.22% increase relative to mean swimmable and fishable, respectively. To put these effects in context, note that from 1996 to 2008 the percentage of swimmable and fishable water bodies increased by 1.2% and 0.4% per year on average, respectively (*SI Appendix*, Table S1). Relative to this underlying rate of change, water group-related increases of 0.38–0.81% in proportion swimmable and 0.22–0.39% in proportion fishable are meaningful. Clearly, the presence of groups did not increase yearly in every watershed. However, these results confirm that water groups had a positive and significant impact on water quality in the watersheds where they are located, as measured by average DOD and proportions of swimmable and fishable water bodies.

Our focus is on measuring causal impacts of water groups on water quality. In this context, whether the results hold across alternative models and data samples is more relevant than goodness of fit. Hence, we check for robustness of our results to a variety of model specifications; see *SI Appendix* and the detailed description

Table 2. Effects of water groups on water quality: Swimmable and fishable

	Swimmable			Fishable		
Explanatory variables	Effect of groups	Effect of donations	Effect of expenditures	Effect of groups	Effect of donations	Effect of expenditures
No. of water groups$_{t-1}$	0.005*** (0.002)			0.003** (0.001)		
Donations$_{t-1}$, 1,000s $		4.5×10^{-07}** (1.8×10^{-07})			3.4×10^{-07}*** (9.7×10^{-08})	
Program expenditures$_{t-1}$, 1,000s $			2.5×10^{-07}*** (6×10^{-08})			1.9×10^{-07}*** (4.1×10^{-08})
Violations$_{t-1}$	−0.179*** (0.051)	−0.181*** (0.050)	−0.181*** (0.051)	−0.061** (0.027)	−0.063** (0.027)	−0.063** (0.027)
Federal water quality	6.5×10^{-07}	6.7×10^{-07}	6.7×10^{-07}	3.2×10^{-07}	3.3×10^{-07}	3.3×10^{-07}
Expenditures, 1,000s $	(6.8×10^{-07})	(6.8×10^{-07})	(6.8×10^{-07})	(3.7×10^{-07})	(3.7×10^{-07})	(3.7×10^{-07})
Fraction agricultural land	−0.067 (0.057)	−0.066 (0.057)	−0.067 (0.057)	−0.062 (0.044)	−0.062 (0.044)	−0.062 (0.044)
Fraction urban land	−0.024 (0.129)	0.008 (0.129)	0.009 (0.129)	−0.092 (0.103)	−0.079 (0.101)	−0.078 (0.101)
Population density, persons/mi^2	-0.9×10^{-04} (0.7×10^{-04})	-0.9×10^{-04} (0.7×10^{-04})	-0.9×10^{-04} (0.7×10^{-04})	-6.9×10^{-05} (5.3×10^{-05})	-7.0×10^{-05} (5.3×10^{-05})	-7.0×10^{-05} (5.3×10^{-05})
Per capita income, 1,000s $	-7.8×10^{-05}*** (1.6×10^{-05})	-7.6×10^{-05}*** (1.6×10^{-05})	-7.6×10^{-05}*** (1.6×10^{-05})	-2.6×10^{-05}*** (7.4×10^{-06})	-2.5×10^{-05}*** (7.2×10^{-06})	-2.5×10^{-05}*** (7.2×10^{-06})
High school education	−0.159 (0.180)	−0.175 (0.177)	−0.171 (0.177)	−0.178 (0.145)	−0.188 (0.145)	−0.185 (0.145)
Unemployment rate	0.067 (0.618)	0.014 (0.617)	0.009 (0.617)	0.081 (0.531)	0.230 (0.531)	0.049 (0.531)
Precipitation, 1,000s mm	9.6×10^{-04} (9.1×10^{-04})	0.001 (9.2×10^{-04})	0.001 (9.3×10^{-04})	3.0×10^{-04} (6.0×10^{-04})	3.7×10^{-04} (6.0×10^{-04})	3.7×10^{-04} (6.0×10^{-04})
Precipitation^2, 1,000s mm^2	-5.3×10^{-06} (3.4×10^{-06})	-5.8×10^{-06}* (3.5×10^{-06})	-5.8×10^{-07}* (3.5×10^{-06})	-1.1×10^{-06} (2.3×10^{-06})	-1.4×10^{-06} (2.3×10^{-06})	-1.4×10^{-06} (2.3×10^{-06})
Percent Republican vote	0.117* (0.069)	0.106 (0.070)	0.106 (0.070)	0.047 (0.051)	0.041 (0.051)	0.041 (0.051)
Home ownership rate	0.425*** (0.120)	0.413*** (0.121)	0.413*** (0.121)	0.408*** (0.098)	0.406*** (0.098)	0.406*** (0.098)
Percent white population	0.370*** (0.110)	0.367*** (0.110)	0.368*** (0.110)	0.196*** (0.068)	0.194*** (0.068)	0.195*** (0.068)
R^2	0.26	0.26	0.26	0.23	0.23	0.23
Observations	7,310	7,310	7,310	7,242	7,242	7,242

Includes year and state fixed effects. *$P < 0.1$; **$P < 0.05$; ***$P < 0.01$. HUC8-clustered SEs are in parentheses.

in *Methods*. Our results hold with alternative fixed-effects structures, the use of lagged variables, the use of a number of alternative water quality data subsamples, and an alternative estimation approach (matching followed by calculation of average treatment effects). Estimated effects of an additional group in a watershed range from 0.0026 to 0.0034 percentage point reductions in DOD; the impact of an additional $100,000 in donations ranges from 0.0024 to 0.0059 percentage point reductions in DOD; and the effect of a $100,000 increase in expenditures ranges from 0.001 to 0.0025 percentage point reductions in DOD.

Discussion

The economics literature on provision of public goods by nonprofits has focused on donations, which are inputs to public good production, rather than on the output of the public good itself. However, little is known about the extent to which donations translate into increased supply of public goods. This article provides a large-scale empirical assessment of the relationship between nonprofit activity and public good provision.

We focus on water quality as a public good because water pollution remains a problem in the United States despite the CWA becoming law over 40 y ago. An Environmental Protection Agency (EPA) report sampled nearly 2,000 locations in 2008 and 2009, from large rivers to small streams, finding over 55% of them in poor biological health (https://www.epa.gov/national-aquatic-resource-surveys/national-rivers-and-streams-assessment). Nonprofit groups are in a unique position to mitigate water pollution. They are knowledgeable about local water quality and are critical in identifying imperiled water bodies. Watershed groups target their efforts through water quality testing, direct cleanups, educational outreach, and interactions with government and local industry. Local residents support these groups through volunteering and donations.

During our study period, as the number of groups increased in a watershed, DOD decreased and the proportion of swimmable and fishable water bodies increased relative to a counterfactual watershed without water groups. Donations to these groups and their expenditures similarly had positive impacts on water quality. While nominally small, these impacts are meaningful relative to underlying water quality trends. Water quality changes slowly and considerable improvements have been achieved,[§] which means there may not be much scope for further large improvements. In addition, the relationship between water group activity and water quality is complicated by hydrological, ecological, and social aspects, with connections and feedbacks in time and space, in what are termed complex adaptive systems (36).

Furthermore, while we focus on the direct impacts of water groups, these groups may also affect water quality indirectly. For instance, many groups interact with local industry and government regulators. Grant and Grooms (35) find that an additional water group in a watershed reduced CWA violations by between 0.1 and 0.16 percentage points. Given estimates for the impact of violations on water quality in our preferred specifications, this translates to an additional group reducing DOD by 3.7–5.9% and increasing the proportion of swimmable and fishable by 2.8–4.2% and 0.8–1.2%, respectively, through this indirect effect.

To circle back to our motivation, a natural question is whether we can measure the value of improvements in water quality and compare them to the costs incurred by watershed groups. Keiser and Shapiro[¶] assess the benefits of water quality improvements by looking at local housing values after a treatment plant receives a grant. They compare the ratio of a grant's effect on housing values to its costs and find that values are less than costs. They do not value the benefits of a stream segment improving in category, fishable/swimmable.

Carson and Mitchell (37) estimate willingness to pay for improvements from boatable to fishable and swimmable for all freshwater bodies in the United States. Boatable is the lowest CWA designated use, defined as water quality that is acceptable for boating. Carson and Mitchell's (37) best estimates for annual household values, adjusted to 2008 dollars, are $154 for maintaining boatable standards, $116 for making all freshwater bodies at least fishable, and $129 for making all water bodies swimmable, for a total willingness to pay of $399. We cannot use these values directly because they are for all water bodies in the entire country, inclusive of lakes and wetlands. However, at a minimum, we know that water quality did not decline on average. Therefore, nonprofit groups provided at least the service of not allowing degradation below boatable. There were 117 million households in the United States in 2008; thus, the aggregate value was $18 billion, well above the total annual expenditures of watershed groups, which are on the order of $5 billion annually for the 48 mainland states.

Houck (38) asserts "citizen groups … have moved many parts of the CWA forward over institutional resistance, and continue to do so." While the impact of nonprofits is in doubt due to inadequate incentives for donors and managers, and competition over donations or authority, our results provide evidence that private groups can indeed mitigate environmental problems. Additional groups are associated with lower DOD, which may demonstrate that they overcome competition for donations and work collectively. It also suggests that nonprofits groups may provide a mechanism to improve coordination to mitigate collective action problems. Moreover, previous literature using monetary donations as the output may actually underestimate the impact: Our results find larger effects accounting for the existence of groups than through their donations and spending. Furthermore, water quality improvements by water groups are not determined by the violation rate of nearby industries. These findings suggest that the private sector promotes compliance with environmental regulations in ways that extend beyond instigating inspections and suing violators. Furthermore, policies that increase contributions to these groups, such as tax deductions for donations to nonprofits, might pay off in terms of complementing government efforts to achieve environmental quality goals. Finally, our results give proof of concept for the private provision of a public good by nonprofit organizations.

Methods

Data. To construct an empirical measure of a public good, it is key to note the appropriate spatial scale. To account for factors affecting water quality, we contend that the ideal size is watershed based. Drainage basin sizes vary and watersheds can be subdivided into smaller areas defined by the forks of main rivers. Thus, we rely on US Geological Survey (USGS) methodology. USGS divides the United States into successively smaller subunits identified by unique hydrologic unit codes (HUCs). USGS names these units with 2–14 digits based on decreasing size class (39). We use the eight-digit HUC (HUC8) designation as the definition of a watershed for two main reasons. First, drainage basins correspond to the natural boundary for surface water flow. Second, nonprofit groups targeting water quality generally operate at a local scale and the HUC8 scale corresponds to the smallest area a single group is likely to affect. There are 2,264 HUC8 watersheds in the United States, averaging 700 mi^2 in land cover size.

We focus on DO to measure water quality. We obtained DO measurements for rivers and streams from the National Water Information System (USGS) and Storet (US EPA) databases (available at the National Water Quality Monitoring Council's Water Quality Portal; https://www.waterqualitydata.us/). The steps followed to construct the water quality sample from these data are described in *SI Appendix*. We use a standard formula to convert DO in milligrams per liter to DO saturation (in percentage), and calculate DOD as 100 – DO (in percent saturation). This process yields 2,276,913 measurements during our study period. Finally, we aggregate by calculating yearly averages of all measurements within each watershed.

We use proportion swimmable and fishable as additional measures of water quality. Swimmable and fishable indicators are often based on several pollutants, namely biochemical oxygen demand, fecal coliforms, and total suspended solids. Since there are fewer measurements for these pollutants than for DO, using them to define swimmable and fishable considerably reduces the size of the sample available for estimation. Hence, we construct indicator variables for these designations based on minimum DO thresholds: 4.99 mg/L for fishable and 6.47 mg/L for swimmable.[#] We calculate the proportion of DO measurements in a watershed that meet each threshold and use these as dependent variables in place of DOD in our models.

We assume that watershed groups provide an appropriate context for assessing the effectiveness of environmental groups because geographical boundaries define a watershed rather than arbitrary political boundaries, and hence oversight by watershed groups takes place within the correct spatial scale to account for relevant factors affecting water quality. Our data on water-focused nonprofit groups come from several sources. An initial search in Guidestar, an organization that gathers information about nonprofits, yielded a comprehensive list of groups working on "Water Resource, Wetlands Conservation, and Management." The information retrieved includes date of incorporation, location, type of group, and the Employer Identification Number (EIN), which is the federal tax identifier. This information does not allow us to further differentiate groups by more specific objectives. We cross-referenced and supplemented this information with lists from EPA and River Network, a national group assisting organizations whose mission is protecting water resources. These two datasets do not contain the EIN number, so we manually searched for and added any missing numbers to our final list. The EIN number links this list to a database from the US Internal Revenue Service with tax return data for each group, including yearly revenues and expenditures.

We estimate the effect of water groups conditional on factors that impact water quality at the watershed level. First, we account for state and federal

[§]Keiser DA, Shapiro JS (2017) Consequences of the Clean Water Act and the demand for water quality (Department of Economics, Iowa State University, Ames, IA). Available at www.nber.org/papers/w23070. Accessed May 6, 2017.

[¶]Keiser DA, Shapiro JS (2017) Consequences of the Clean Water Act and the demand for water quality (Department of Economics, Iowa State University, Ames, IA). Available at www.nber.org/papers/w23070. Accessed May 6, 2017.

[#]Smith VK, Wolloh CV (2012) Has surface water quality improved since the Clean Water Act? (NBER, Cambridge, MA). Available at www.nber.org/papers/w18192. Accessed November 28, 2017.

enforcement of discharges covered under the CWA by including the total number of discharge permit violations in a watershed in each year. Information on violations and location of facilities are from the EPA's Enforcement and Compliance History Online database.

We account for federal programs that aim to improve water quality by including total expenditures in the state where a watershed is located in each year under three key programs: the Conservation Reserve Program (CRP), the EPA 319 Grant Program, and the Environmental Quality Incentives Program (EQUIP). Data on CRP payments were obtained from the US Department of Agriculture Farm Service Agency (https://www.fsa.usda.gov/programs-and-services/conservation-programs/reports-and-statistics/conservation-reserve-program-statistics/index), data on payments made under the EPA 319 program are from the Grants Reporting and Tracking System (https://iaspub.epa.gov/apex/grts/f?p=109:9118), and information on payments made under EQUIP contracts was provided by the Environmental Working Group (https://www.ewg.org/).

Land use is another factor affecting water quality. Agriculture is a leading cause of impairment of surface waters (40). In urban areas, impervious surfaces contribute to increased pollutant loads and surface runoff of contaminants and sediment, as well as larger variances in stream flow and temperature (41). We account for these effects by using land cover maps from the Multi-Resolution Land Characterization Consortium to calculate proportions of urban and agricultural land in each watershed for each year.

Precipitation is another determinant of water quality. Small amounts of precipitation wash pollutants into water bodies through runoff, whereas relatively large amounts of precipitation accelerate dilution of pollutants. We include mean precipitation in a watershed and its square for each year using data from the PRISM Climate Group (42).

We condition on environmental and other political preferences of the residents of a watershed, as they may affect political outcomes or other factors that can affect water quality. To the extent that these preferences are expressed through voting, they can be accounted for by using election outcomes. We use the proportion of votes for Republican candidates in US Senate races using county-level data on election results from the CQ Press Voting and Elections Collection, interpolating for years in which there were no Senate races, and reweighting to the HUC8 level based on the proportions of counties contained in a watershed.

Finally, demographic characteristics affect water quality. We include population density, per capita income (in 2008 dollars), percentage of population with a high school degree, percentage of the population that are white, home ownership rate, and unemployment rate. This information is at the county level from the US Census and the Bureau of Economic Analysis. Population, income, and unemployment are available for every year in our study. Ethnicity, educational attainment, and home ownership are available for 1990, 2000, and 2010, so we interpolate for intracensus years. We reweight and aggregate these data to the watershed level.

We construct a panel dataset of these variables for 1,131 HUC8 watersheds in the 48 mainland states during 1995–2008. Our study period corresponds to years with reliable data on water groups. Summary statistics (*SI Appendix*, Table S1) suggest that water quality and water group activity, and indeed most other watershed characteristics, change gradually over the study period, reflecting slow-moving underlying processes.

Empirical Strategy. An ideal research design would use data from an experiment in which water groups are randomly placed across watersheds. This, of course, is not feasible, and in practice the main challenge for measuring causal effects of water group activity on water quality is that water groups do not locate randomly across watersheds. Indeed, these groups may locate where water quality is relatively poor, which may lead to more donations and higher expenditures. As a result, watershed group activity and water quality may be jointly determined, and ordinary least-squares estimates may not be consistent. To address this concern, we follow a growing number of studies that combine matching and panel methods, particularly fixed-effects estimation (43–50).

We define treated watersheds as having at least one water-focused group during the study period, and control watersheds as those that do not have any groups during the entire study period. We preprocess the data to make treated and control watersheds observationally similar before the study period by matching on time-invariant or pretreatment observable characteristics that affect water quality: 1995 values for CWA violations, federal programs, precipitation, proportions of urban and rural land, proportion of vote for Republican candidates, per capita income, population density, high school graduation, ethnicity, unemployment, and home ownership rate. Additionally, we match within state because states are responsible for impairment listing of water bodies in the watershed. This accounts for state-level characteristics that may have an impact on water quality. [This accounts for the EPA requiring each state to submit a list of all threatened and impaired waters during even-numbered years (51) and provides information on which state is responsible for listing when water bodies cross state lines.] Finally, we match on water quality (DOD) in 1995. This helps mitigate the concern that water-focused groups may locate in watersheds with poorer water quality, since treated and control watersheds used in the estimation sample have similar DOD measures at the beginning of the study period. More details on the matching procedure and the balance of the estimation sample are presented in *SI Appendix*.

For the postmatching balanced sample, our regression models for mean DOD, proportion swimmable, and proportion fishable in watershed i in year t are as follows:

$$\mathrm{DOD}_{it} = \alpha_1 \mathrm{Group\ Activity}_{it-1} + \mathrm{CWA}_{it-1}\alpha_2 + \boldsymbol{X}_{it}\boldsymbol{\alpha}_3 + \boldsymbol{\delta}_s + \boldsymbol{\tau}_t + \varepsilon_{it},$$

$$\mathrm{Swimmable}_{it} = \beta_1 \mathrm{Group\ Activity}_{it-1} + \mathrm{CWA}_{it-1}\beta_2 + \boldsymbol{X}_{it}\boldsymbol{\beta}_3 + \boldsymbol{\phi}_s + \boldsymbol{\theta}_t + u_{it},$$

$$\mathrm{Fishable}_{it} = \gamma_1 \mathrm{Group\ Activity}_{it-1} + \mathrm{CWA}_{it-1}\gamma_2 + \boldsymbol{X}_{it}\boldsymbol{\gamma}_3 + \boldsymbol{\psi}_s + \boldsymbol{\omega}_t + v_{it}.$$

Group Activity$_{it}$ is measured three ways: total number of groups, total donations, and total expenditures in watershed i and year t. Donations and expenditures are expressed in 2008 dollars. We include lagged effects of water group activity and CWA permit violations because watershed impairment can be a slow-moving process, and these factors likely have an impact with a lag. The matrix $\boldsymbol{X}_{it}$ contains contemporaneous values of the explanatory variables discussed above. The vectors $\boldsymbol{\tau}_t$, $\boldsymbol{\theta}_t$, and $\boldsymbol{\omega}_t$ are year fixed effects. The vectors $\boldsymbol{\delta}_s$, $\boldsymbol{\phi}_s$, and $\boldsymbol{\psi}_s$ are fixed effects for the state responsible for impairment listing of water bodies in the watershed. This is the appropriate level to introduce fixed effects because water body management and impairment listing decisions are made annually at the state level. We avoid using watershed fixed effects because they remove the signal we wish to measure in the data (for example, see www.g-feed.com/2012/12/the-good-and-bad-of-fixed-effects.html) and thus eliminate most power of our explanatory variables. Fixed effects are preferred to random effects because the former allow us to control for unobserved state-level heterogeneity that is assumed to be correlated with included explanatory variables. Random effects, in contrast, are assumed to be independent of included regressors.

Alternative Specifications. Our main models use fixed effects for the state responsible for impairment listing of water bodies. We consider two additional fixed-effects structures. First, we use state–year fixed effects to allow for more flexible time-variant state heterogeneity. Second, we use fixed effects for combinations of states included in transboundary watersheds to account for watersheds that extend across state lines and hence are subject to different jurisdictions. The results (*SI Appendix*, Tables S3 and S4) are consistent with those from our preferred specifications, except for the coefficient for number of groups, which is not statistically significant when using transboundary fixed effects. We also estimated the models without lagged regressors and with all regressors lagged. The results are consistent with those presented in Table 1. Finally, we estimated the models without the conditioning variables. If our approach for addressing nonrandom water group activity is effective, the water group variables should have the same sign as in the fully specified models. The results show that this is the case (*SI Appendix*, Table S5).

There is considerable spatial and temporal variation in water quality monitoring, and local governments or private organizations commonly choose the location and frequency of sampling. Some states have denser monitoring networks, waters near populated areas tend to be overrepresented, and monitors drop in and out over time (52).[‖,**] We conduct several sensitivity tests to ensure that our results are not driven by these characteristics of the DOD data. First, we use only monitors that have at least 25 measurements, as these may have higher-quality data. Second, we only use measurements from well-documented and high-quality monitors from the National Water Quality Assessment network.** Results (*SI Appendix*, Tables S6 and S7) are consistent with those from the full dataset. We also use summer rather than yearly DOD averages. Additionally, 10 states did not consistently file impairment reports during the study period (Utah, Vermont, Pennsylvania, New York, Nebraska, Mississippi, Minnesota, Connecticut,

‖Smith VK, Wolloh CV (2012) Has surface water quality improved since the Clean Water Act? (NBER, Cambridge, MA). Available at www.nber.org/papers/w18192. Accessed November 28, 2017.

**Keiser DA, Shapiro JS (2017) Consequences of the Clean Water Act and the demand for water quality (Department of Economics, Iowa State University, Ames, IA). Available at www.nber.org/papers/w23070. Accessed May 6, 2017.

Colorado, and Washington). If this is due to unreliable monitoring, the data from these states may be relatively poor, so we run our models without these states. Results for these sensitivity checks (*SI Appendix*, Tables S8 and S9) are consistent with those of our preferred specifications.

Finally, we assess the robustness of our results to a different estimation approach. We define the outcome of interest as the percentage change in DOD between 1996 and 2008, and consider the same treatment (presence of at least one watershed group during the study period). Instead of matching followed by fixed-effects regression, we use matching (with four nearest neighbors) and then nonparametrically estimate the average treatment effect (ATE). This approach hinges on the conditional independence (unconfoundedness) assumption that, conditional on observed covariates, treatment is independent of the outcome (details are available in *SI Appendix*). We check for robustness to alternative numbers of nearest neighbors (1–3, 5). We also assess sensitivity to omitting matching covariates and conduct a falsification test by estimating the ATE on a pretreatment outcome (change in DOD between 1992 and 1995). Results are presented in *SI Appendix*, Table S10.

The results support those from our main models. Water groups had a positive and significant impact on water quality, as the rate increase of DOD is smaller by 1.8 percentage points in watersheds where water groups were located. This effect is robust to the number of nearest neighbors used to construct the counterfactual matches. Additionally, the results are robust to the exclusion of covariates used for matching. Finally, there is no effect on the pretreatment outcome, which suggests there is no selection on unobserved factors (53).

Our paper gives proof of concept for these watershed groups providing the public good of improved water quality. Stronger results could be obtained with additional data: (*i*) an annual survey of watershed groups project regarding water quality, in other words, better expenditure information including where and how the money is spent, or (*ii*) randomizing grant funding to these groups, or (*iii*) better sampling of water quality (more temporally consistent and more spatially representative). These data do not exist, but their collection would further this research in the area of private provision of public goods by environmental groups.

1. Levin SA (2014) Public goods in relation to competition, cooperation, and spite. *Proc Natl Acad Sci USA* 111:10838–10845.
2. The Center on Philanthropy at Indiana University (2017) *Giving USA 2017. The Annual Report on Philanthropy for the Year 2016* (Indiana University, Indianapolis, IN).
3. Sugden R (1984) Reciprocity: The supply of public goods through voluntary contributions. *Econ J* 94:772–787.
4. Roberts RD (1987) Financing public goods. *J Polit Econ* 95:420–437.
5. Bergstrom TC, Blume L, Varian H (1986) On the private provision of public goods. *J Public Econ* 33:25–49.
6. Andreoni J (1989) Giving with impure altruism: Applications to charity and Ricardian equivalence. *J Polit Econ* 97:1447–1458.
7. Holländer H (1990) A social exchange approach to voluntary cooperation. *Am Econ Rev* 80:1157–1167.
8. Glazer A, Konrad KA (1996) A signaling explanation for charity. *Am Econ Rev* 86: 1019–1028.
9. Harbaugh WT (1998) What do donations buy? A model of philanthropy based on prestige and warm glow. *J Public Econ* 67:269–284.
10. Ribar DC, Wilhelm MO (2002) Altruistic and joy-of-giving motivations in charitable behavior. *J Polit Econ* 110:425–457.
11. Kingma RB (1989) An accurate measurement of the crowd-out effect, income effect, and price effect for charitable contributions. *J Polit Econ* 97:1197–1207.
12. Posnett J, Sandler T (1989) Demand for charity donations in private non-profit markets: The case of the U.K. *J Public Econ* 40:187–200.
13. Smith VH, Kehoe MR, Cremer ME (1995) The private provision of public goods: Altruism and voluntary giving. *J Public Econ* 58:107–126.
14. Payne A (1998) Does the government crowd-out private donations? New evidence from a sample of non-profit firms. *J Public Econ* 69:323–345.
15. Okten C, Weisbrod BA (2000) Determinants of donations in private nonprofit markets. *J Public Econ* 75:255–272.
16. Kotchen MJ, Moore MR (2007) Private provision of environmental public goods: Household participation in green-electricity programs. *J Environ Econ Manage* 53: 1–16.
17. Andreoni J, Payne A (2003) Do government grants to private charities crowd out giving or fundraising? *Am Econ Rev* 93:792–812.
18. Duquette NJ (2016) Do tax incentives affect charitable contributions? Evidence from public charities' reported revenues. *J Public Econ* 137:51–69.
19. Boberg-Fazlić N, Sharp P (2017) Does welfare spending crowd out charitable activity? Evidence from historical England under the poor laws. *Econ J* 127:50–83.
20. Andreoni J, Payne A (2011) Is crowding out due entirely to fundraising? Evidence from a panel of charities. *J Public Econ* 95:334–343.
21. Duncan B (1999) Modeling charitable contributions of time and money. *J Public Econ* 72:213–242.
22. Rose-Ackerman S (1996) Altruism, nonprofits, and economic theory. *J Econ Lit* 34: 701–728.
23. Glaeser EL (2003) Introduction. *The Governance of Not-for-Profit Organizations*, ed Glaeser EL (The University of Chicago Press, Chicago).
24. Castaneda MA, Garen J, Thornton J (2007) Competition, contractibility, and the market for donors to nonprofits. *J Law Econ Organ* 24:215–246.
25. Werker E, Ahmed FZ (2008) What do nongovernmental organizations do? *J Econ Perspect* 22:73–92.
26. Rose-Ackerman S (1982) Charitable giving and excessive fundraising. *Q J Econ* 97: 193–212.
27. Aldashev G, Verdier T (2010) Goodwill bazaar: NGO competition and giving to development. *J Dev Econ* 91:48–63.
28. Nunnenkamp P, Öhler H (2012) Funding, competition, and the efficiency of NGOs: An empirical analysis of non-charitable expenditure of US NGOs engaged in foreign aid. *Kyklos* 65:81–110.
29. Henderson JV, Lee YS (2015) Organization of disaster aid delivery: Spending your donations. *Econ Dev Cult Change* 63:617–664.
30. Baik KH, Shogren JF (1994) Environmental conflicts with reimbursement for citizen suits. *J Environ Econ Manage* 27:1–20.
31. Heyes AG (1997) Environmental regulation by private contest. *J Public Econ* 63: 407–428.
32. Sundberg JO (2006) Private provision of a public good: Land trust membership. *Land Econ* 82:353–366.
33. Albers HJ, Ando AW, Chen X (2008) A spatial-econometric analysis of attraction and repulsion of private conservation by public reserves. *J Environ Econ Manage* 56:33–49.
34. Langpap C, Shimshack J (2010) Private citizen suits and public enforcement: Substitutes or complements? *J Environ Econ Manage* 59:235–249.
35. Grant LE, Grooms K (2017) Do nonprofits encourage compliance? *J Assoc Environ Resour Econ* 4:S261–S288.
36. Lubchenco J, Cerny-Chipman EB, Reimer JN, Levin SA (2016) The right incentives enable ocean sustainability successes and provide hope for the future. *Proc Natl Acad Sci USA* 113:14507–14514.
37. Carson R, Mitchell R (1993) The value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resour Res* 29: 2445–2454.
38. Houck OA (2014) Cooperative federalism, nutrients, and the Clean Water Act: Three cases revisited. *Environ Law Report* 44:10426–10442.
39. Seaber PR, Kapinos FP, Knapp GL (1987) Hydrologic unit maps. US Geological Survey Water Supply Paper 2294(US Geological Survey, Reston, VA). Available at https://pubs.usgs.gov/wsp/wsp2294/pdf/wsp_2294_a.pdf. Accessed April 3, 2016.
40. Baker LA (1992) Introduction to nonpoint source pollution in the U.S. and prospects for wetland use. *Ecol Eng* 1:1–26.
41. Tang Z, Engel BA, Pijanowski BC, Lim KJ (2005) Forecasting land use change and its environmental impact at a watershed scale. *J Environ Manage* 76:35–45.
42. PRISM Climate Group Oregon State University (2018) PRISM Climate Data. Available at prism.oregonstate.edu. Accessed May 14, 2017.
43. Imbens GW, Wooldridge JM (2009) Recent developments in the econometrics of program evaluation. *J Econ Lit* 47:5–86.
44. Arriagada RA, Ferraro PJ, Sills EO, Pattanayak SK, Cordero-Sancho S (2012) Do payments for environmental services affect forest cover? A farm-level evaluation from Costa Rica. *Land Econ* 88:382–399.
45. Ferraro PJ, Miranda JJ (2014) The performance of non-experimental designs in the evaluation of environmental policy: A design-replication study using a large-scale randomized experiment as a benchmark. *J Econ Behav Organ* 107:344–365.
46. Ferraro PJ, Miranda JJ (2017) Panel data designs and estimators as substitutes for randomized controlled trials in the evaluation of public programs. *J Assoc Environ Resour Econ* 4:281–317.
47. Alix-Garcia JM, Sims KRE, Yañez-Pagans P (2015) Only one tree from each seed? Environmental effectiveness and poverty alleviation in Mexico's payments for ecosystem services program. *Am Econ J Econ Policy* 7:1–40.
48. Jones KW, Lewis DJ (2015) Estimating the counterfactual impact of conservation programs on land cover outcomes: The role of matching and panel regression techniques. *PLoS One* 10:e0141380.
49. Wendland KJ, Baumann M, Lewis DJ, Sieber A, Radeloff VC (2015) Protected area effectiveness in European Russia: A post-matching panel data analysis. *Land Econ* 91: 149–168.
50. Chen Y, Lewis DJ, Weber B (2016) Conservation land amenities and regional economies: A post-matching difference-in-differences analysis of the Northwest Forest Plan. *J Reg Sci* 56:373–394.
51. US Environmental Protection Agency (2009) Fact Sheet: Introduction to Clean Water Act Section 303(d) Impaired Waters Lists. Available at https://cfpub.epa.gov/watertrain/pdf/modules/introtocwa.pdf. Accessed September 15, 2018.
52. Sigman H (2005) Transboundary spillovers and decentralization of environmental policies. *J Environ Econ Manage* 50:82–101.
53. Imbens GW, Rubin D (2015) *Causal Inference for Statistics, Social and Biomedical Sciences* (Cambridge Univ Press, Cambridge, UK).