Community-based monitoring to facilitate water management by local institutions in Costa Rica

María Bernedo Del Carpio[a,1], **Francisco Alpizar**[b,c], **and Paul J. Ferraro**[d,e]

[a]Department of Economics, University of Maryland Baltimore County, Baltimore, MD 21250; [b]Department of Social Sciences, Wageningen University and Research, 6700 EW Wageningen, The Netherlands; [c]Environment for Development Program, Tropical Agricultural Research and Higher Education Center, Turrialba, Costa Rica, 30501; [d]Carey Business School, Johns Hopkins University, Baltimore, MD 21202; and [e]Department of Environmental Health and Engineering, a joint department of the Bloomberg School of Public Health and the Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21212

Edited by Arun Agrawal, University of Michigan, Ann Arbor, MI, and approved March 29, 2021 (received for review August 3, 2020)

Water scarcity is a global problem that can be compounded by inefficient water management, including underinvestment in infrastructure, underpricing of water use, and underenforcement of user rules. Here, we explore whether these inefficiencies can be reduced in rural Costa Rica via an externally driven community monitoring program (i.e., a program initiated by an outside organization and run by citizens). The monitoring program aimed to reduce groundwater extraction from aquifers, as well as to improve water quality and user satisfaction, by supplying additional information about field conditions and additional scrutiny of user and management authority activities and by fostering citizen engagement in water management. Using a specially designed smartphone application (app) and WhatsApp, monitors could report weekly on the conditions of the water system, including service disruptions, water quality, leaks, and source contamination. The app automatically compiled the individual reports into a summary report, which was then made available to the community water management committees and water users. The program was randomly implemented in 80 of 161 communities that expressed an interest in participating. One year after the program started, we detect modest, albeit imprecisely estimated, effects of the program in the predicted directions: less groundwater extracted, better water quality, and more satisfied users. Although the estimated effects are imprecise, the monitoring program appears to be equally or more cost effective for reducing groundwater extraction than another program in the same region that encouraged households to adopt water-efficient technologies.

community-based natural resource management | community-based environmental monitoring | common pool resource | collective action | citizen engagement

Since the 1980s, a growing group of scholars and practitioners has advocated for the decentralization of natural resource governance. In their view, the participation of local communities in resource governance leads to better environmental and social outcomes (1–3). Here, we focus on one aspect of decentralized natural resource governance: community-based monitoring, in which organized groups of citizens collect information on the state of a resource and resource user activities in their communities and share this information with other citizens and resource management authorities.

The potential channels through which community-based monitoring can improve resource management outcomes are highlighted in three overlapping literatures: the literature on common pool resources (1), the literature on community-based environmental monitoring [which overlaps with the literature on citizen science (4, 5)], and the literature on citizen monitoring of public services (6). These literatures highlight how monitoring systems can improve resource management by providing new information about a resource and the behaviors of users, by fostering more citizen engagement in resource management, and by holding users and management authorities more accountable for their behaviors.

The literature on common pool resources focuses on community-based monitoring as a means through which unauthorized behaviors can be materially sanctioned or through which social norms can be leveraged (often in combination with communication among users). In this literature, an association between community-based monitoring and better environmental quality has been described in theoretical frameworks and suggested by numerous case studies (e.g., refs. 7–9). Yet establishing a causal connection (10) between changes in monitoring and improved resource outcomes is challenging in case studies (11). Even in large-sample observational studies in which monitoring has arisen endogenously in some communities and not in others, scholars cannot easily disentangle the effects of monitoring separate from the effects of other community attributes that are correlated with both monitoring and resource conditions. The endogenous nature of these monitoring systems not only affects the internal validity of the study designs (i.e., whether monitoring truly causes better environmental conditions), but it also affects their generalizability to contexts in which community-based monitoring systems do not yet exist or are weak and where external actors are encouraging their creation—what we call externally driven, community-based monitoring. These contexts are common, particularly in efforts to decentralize natural resource management authority to communities. Externally driven monitoring is common in laboratory experiments, which randomize monitoring to isolate its causal effects (e.g., refs. 12–15). Yet whether the insights from

Significance

Elinor Ostrom won the Nobel Prize in Economics for demonstrating that humans can create rules and institutions that permit sustainable management of shared resources without resorting to privatization or government expropriation. One purported enabling condition for success is monitoring of the shared resource by community members. Whether such monitoring can be encouraged where it is absent, and thereby improve resource management, is not well understood. In a randomized trial, we assessed whether an externally encouraged, community-based monitoring program improved water management. After 1 y, we detect modest reductions in groundwater pumping and modest improvements in water quality and user satisfaction. Although replications are needed, the results imply that externally encouraged, community-based monitoring can improve the management of shared resources.

Author contributions: M.B.D.C., F.A., and P.J.F. designed research; M.B.D.C. and F.A. performed research; M.B.D.C. and P.J.F. analyzed data; and M.B.D.C., F.A., and P.J.F. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: mbernedo@umbc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2015177118/-/DCSupplemental.

https://doi.org/10.1073/pnas.2015177118

these laboratory settings generalize to field settings is unclear. Moreover, the experiments typically randomize treatments that combine monitoring and sanctioning, and thus the individual contribution of adding or strengthening just the monitoring system remains unclear.

In the literature on community environmental monitoring, scholars focus on monitoring as a means through which resource management-relevant information can be collected by the people in the best position to collect and disseminate it (i.e., people who live and work in the monitored environment). By disseminating this information to other citizens and management authorities, monitors contribute to better management decisions (16, 17). The literature also posits that community monitoring fosters engagement and more active participation by citizens in resource management, which can also improve management decisions (17). The citizen monitoring programs in this literature are typically externally encouraged (and sometimes externally managed), and thus may offer insights into the impacts of monitoring in programs that decentralize resource management authority. Yet the causal effects of the monitoring programs on environmental conditions are unclear. Although there are cases in which monitoring data has been used in conservation actions (16, 18), we know of no studies that provide empirical evidence that community-based monitoring improved environmental outcomes (4, 19).

The literature on the monitoring of public services and on community involvement in public service delivery focuses on citizen scrutiny as a mechanism to hold management authorities accountable, whether they be local, regional, or national authorities (6, 20). In this literature, externally driven monitoring programs have been experimentally tested in naturally occurring field settings. However, these field experiments have not been conducted in natural resource contexts, but rather in contexts like health, transportation, and education. Moreover, these experimental studies have yielded mixed results on the impacts of the monitoring programs (21–24).

Thus, whether externally driven, community-based monitoring of natural resources yields positive outcomes is an open empirical question. This question is also policy relevant, given that many communities that manage common pool resources do not have fully functioning monitoring systems and may benefit from external encouragement and support for such systems. To contribute to answering this question, we focus on freshwater resources. Freshwater scarcity is an important constraint to sustainable development, with around 70% of the world's population experiencing moderate to severe water scarcity at least 1 mo per year (25). Globally, groundwater provides around one-third of humans' freshwater requirements (26) but is only slowly renewed by precipitation, making improved groundwater management critical. Water scarcity arises not only from geophysical conditions, but also from economic and institutional conditions. For example, inadequate infrastructure and poor management, particularly in low- and middle-income nations, exacerbate geophysical conditions like low precipitation and slow aquifer recharge. Improvements in these economic and institutional conditions can thus help mitigate the effects of water scarcity (27).

To improve these conditions in western Costa Rica, where communities rely on pumping water from aquifers and the resource is managed by community organizations, a nongovernmental organization created a community-based monitoring program. The program created a system of monitors who could provide their communities with additional information about water quality, water continuity, leaks, source contamination, and unsanctioned uses in the community. The monitors gathered and reported the information using a specially designed cell-phone application (app). The program was randomly assigned among 161 communities that expressed an interest in implementing the program. We leverage that random assignment to evaluate the program's impacts on resource use, resource quality, and user satisfaction and to elucidate the channels through which the program affected behaviors within the communities.

Community-Based Water Management Organizations

In high-income countries or in urban areas of low- or middle-income countries, governments and regulated utilities typically supply water to residents and businesses. In contrast, community-based water management organizations (CBWMOs) are important suppliers of water in rural communities in Latin American and Caribbean countries as well as in low- and middle-income countries elsewhere (28–30). CBWMOs are managed by committees of elected citizens who are involved in everything from billing to maintenance of infrastructure to protection of watersheds. CBWMOs are often regulated by governments but typically receive little or no government support. Their lack of external support and their location in poor rural communities with little capacity for investment negatively affect their ability to deliver high-quality water services and adapt to changes in water scarcity.

In Costa Rica, around 1,400 CBWMOs serve ~1.5 million people or nearly 30%of Costa Rica's population (31). Our study communities are mostly located along the Pacific coast and the northern plains—regions that experience a pronounced dry season and are forecasted to experience a further 20% decrease in water availability by 2050 due to climate change (32). Exacerbating these unfavorable geophysical conditions are water management inefficiencies. CBWMOs frequently lack the financial and technical resources to maintain their infrastructure properly. For example, the Costa Rican Institute of Water and Sanitation (AyA), which oversees the provision of drinking water, estimates that around 50% of the water pumped from aquifers is lost in transmission from leaks (33, 34). Management inefficiencies affect not only water quantity: Many CBWMOs do not chlorinate their water, do not test their water quality, and do not protect their water sources from contamination.

Around 70% of the study communities are "formal" CBWMOs. The main difference between formal and informal CBWMOs is that the formal CBWMOs have signed a delegation agreement with AyA, which formally translates into higher accountability to AyA. However, prior research reports that the accountability to AyA is nearly absent, and the formal–informal legal difference does not affect CBWMO performance (35). In practice, the management committees of formal and informal CBWMOs have total control over the management of the water system.

The Monitoring Program

Prior to the program's implementation, the monitoring of the resource's status and use in the study communities was formally done by the members of the CBWMO management committees (third-party monitoring) and informally done by community members (mutual monitoring). The monitoring program, implemented by the Tropical Agricultural Research and Higher Education Center (CATIE), aimed to formalize community member monitoring by designating a small set of volunteers to serve as formal monitors (i.e., adding another third-party monitor). The CATIE monitoring program was part of a larger Evidence in Governance and Politics Metaketa initiative of six coordinated field experiments that tested how external support for community monitoring affects the overuse or degradation of resources (36).

To recruit communities to the monitoring program, the project team created a list of CBWMOs that satisfied the following criteria: 1) the community only pumped water from underground sources, 2) the community's electricity records for its water pump were available from the electricity service provider, 3) the community had internet service, and 4) the CBWMO management committee expressed an interest in participating in the monitoring program. A total of 161 CBWMOs met all the eligibility criteria. *SI Appendix*, Fig. S1 shows their locations.

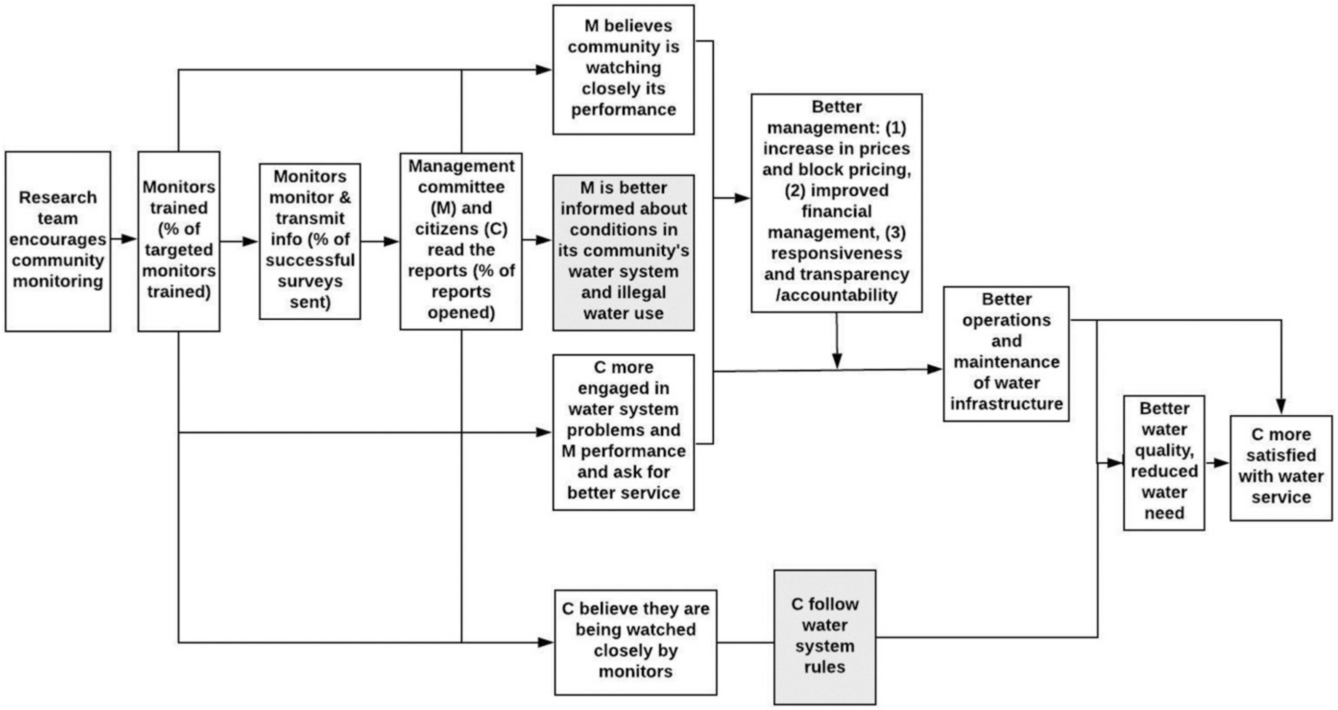

Fig. 1. Theory of change. The figure shows the mechanisms through which the monitoring program is hypothesized to affect water pumped, water quality, and user satisfaction. We measured variables associated with the boxes in white.

Using block randomization, 80 CBWMOs were randomly assigned to the program group and 81 to the control group (*SI Appendix*). The first step in establishing the monitoring program was a public workshop convened by the CBWMO committee and facilitated by the project team. All residents, including members of the CBWMO committees, were invited to participate. In the workshop, the project team presented the monitoring system and the communication technology, which operated through a smartphone monitoring app called SIMA (Sistema de Monitoreo del Agua) and a WhatsApp chat group.* The project team explained how to install SIMA and helped attendees to download it. The team also taught attendees how to read the summary report that the app generates from the individual monitor reports. To encourage attendees to volunteer as monitors and comply with the weekly reporting targets, the project team explained that monitors would contribute to improving the water system in their community (a public good) and that they would receive a payment of about US\$3[†] for each report submitted (a private good), paid by the project team. These payments were designed to achieve two objectives: 1) to mitigate the free-riding problem in community-based monitoring programs (i.e., a second-order social dilemma in the management of common pool resources, see ref. 37)[‡] and 2) to compensate monitors for potential increases (perceived or actual) in their phone and internet bills resulting from using the app. A third, unintended benefit from externally controlled payments for monitors was to lower the risk of co-option ("capture") of the monitors by the management committee (a risk that was also reduced by the framing of the monitoring program as an information-sharing initiative to help the management committee rather than as an accountability initiative). The total number of monitors was limited by the project budget. Each community had a predetermined target number of monitors, which was based on the number of connections reported by the CBWMO (Table 1).

To serve as monitors, volunteers needed to satisfy the following criteria: be literate, live in the community, have a smartphone, be willing to serve as a monitor for 9 mo, have internet access, and work neither in nor for the management committee. Volunteers received one-on-one training on how to use the SIMA app. The app, which was developed for the project, allowed monitors to submit information during a weekly reporting window (Friday to Sunday). Using a survey interface on the app, monitors reported what they saw during the week regarding the following: 1) the days when water service was disrupted, 2) the maximum number of hours without service, 3) the days when water did not run clear, 4) the days when water presented an unpleasant or unusual taste or smell, 5) the number and location of new leaks in the pipelines, 6) the number and location of old leaks that had not yet been fixed, 7) problems related to illegal use (e.g., unauthorized water uses or water connections), and 8) unauthorized use of the land around the water source. Every Monday, the app used the information provided by the monitors to create a weekly summary report that all app users could read.

The project team also created a WhatsApp chat group, to which the team added the phone numbers of committee members, monitors, and the project team. The chat group allowed for immediate and less structured communication among monitors, between monitors and CBWMO committee members, and among the project team, monitors, and committee members. There was no monetary incentive to use the WhatsApp chat.

Theory of Change

Fig. 1 illustrates the preregistered theory of change that underlies the intervention (38). Building on theories from the three literatures described in the introduction, the project team hypothesized

---

*The use of smartphones is common in Costa Rica. In a 2017 national survey conducted by the Costa Rican government, 94% of the sample had a cell phone and 87% had a smartphone (44).

[†] The minimum daily wage for nonqualified workers in Costa Rica in 2018 was ¢10,060.75 or US\$17.80. The monitor payment per report sent is around one-sixth of the minimum daily wage (45).

[‡] Monitors were paid in all programs in the six-country Metaketa.

Table 1. Predetermined number of monitors per treated CBWMO

Minimum number of connections per CBWMO	Maximum number of connections per CBWMO	Number of CBWMOs	Number of monitors per CBWMO	Total number of monitors
17	300	63	3	189
301	400	5	4	20
401	500	5	5	25
501	780	7	6	42
Total				276

that the monitoring system could create greater community interest in better management of the water system, create greater accountability by creating a greater sense among users and committees that their actions are being closely watched, and put better information in the hands of the committees regarding the real-time conditions of their water supply. Changes in these mechanisms could in turn lead to better maintenance of infrastructure (repair leaks, reduce overflows, hire a plumber at least part time, test and improve water quality), better management procedures [pricing, financial management, and transparency that conforms with best practice guidelines for CBWMO (39)], and fewer violations of water use rules (fewer informal connections, fewer commercial uses, protection of water source). Better maintenance of infrastructure could improve water quality and reduce water waste, which could reduce groundwater pumping and improve the cost effectiveness of water delivery. Better management procedures could directly affect water consumption by increasing prices to maximum permissible increasing block levels, or it could lead to more investment in maintenance and infrastructure by improving financial management. Fewer violations of water use rules could improve water quality, reduce water consumption, and increase CBWMO income, which in turn could increase investment in the water system. Separate from its impacts on the common pool resource, the monitoring program was also expected to improve how customers perceive their CBWMO committees, thereby improving their overall satisfaction with the water service.

An underlying assumption of our theory of change is that CBWMO do not face unsurmountable obstacles to adjust their management in reaction to increased monitoring and information. Informed by previous studies and policy recommendations (35), we believe that assumption is credible: management committees already have the institutional structure and mandate to invest, manage staff, and set prices in reaction to information.

With this theory of change, we do not aim to produce a complete theory of behavior that fully describes all the very complex interactions that link community monitoring to resource management. Our aim is to provide a conceptual model that identifies the key mechanisms behind a hypothesized causal pathway.

Hypotheses

Informed by the theory of change, we test three preregistered main hypotheses:

H1. Community monitoring reduces the quantity of water pumped from aquifer. Most CBWMOs in the region do not have water meters on their groundwater pumps. To detect changes in pumping, we use monthly metered electricity consumption by the pumps (kWh). We assume that changes in monthly electricity consumption are directly correlated with changes in monthly water pumped (we assess this assumption in the section Cost-Effectiveness Analysis).

H2. Community monitoring improves water quality. To measure water quality in samples taken from each community, the project team used a professional testing kit. The primary outcome measure is the presence of bacteria (binary variable).

H3. Community monitoring increases users' satisfaction with the resource, management, and community use. To measure user satisfaction, we use household survey data to construct two outcome measures: 1) a user satisfaction index that takes on values between 1 and 5 and is derived from five variables that measure people's opinions about water continuity, water quantity, water pressure, water quality, and the work of the CBWMO committee and 2) a measure of users' belief that their community wastes water (binary variable).

To further evaluate the plausibility of the theory of change, we also test the program's effect on twelve intermediate outcome variables. The variables used in this exploratory analysis and the mechanisms for which they serve as a proxy are listed in Table 2 (*SI Appendix* for more details). Implicit in the theory of change are several process steps: the project team had to recruit and train monitors, the monitors had to monitor and send their individual reports, and committee members had to read the summary reports. We measure and report compliance at each step using data from the workshops and the app. To measure the variable "read report," we assume that a person that opens the summary report reads it. A monitoring system that experienced all three steps was labeled a "working system" in the preanalysis plan (PAP). Note that if the latter two steps failed to materialize, the program could still affect the ultimate outcomes if the public workshops in which monitors were trained had a direct effect on committee member actions (e.g., by changing norms or expectations). However, we would be unable to quantify the relative contribution of this pathway unless monitors failed to submit any reports (i.e., compliance with the reporting was 0%).

Results

Compliance. In 69 out of the 80 CBWMOs selected for treatment (86.3%), the project team was able to conduct a workshop and train at least one monitor. Noncompliance resulted from conflicts among the management committee members (three cases), changes in level of interest or time available from the management committee members (six cases), and institutional changes (two CBWMOs were absorbed by another CBWMO).

Fig. 2 shows the process variables. Fig. 2*A* reports the fraction of monitors trained per CBWMO measured as the number of monitors trained divided by the number of monitors assigned to a CBWMO. In 80% of the treated CBWMOs, the field team trained the total number of monitors assigned to the CBWMOs. Fig. 2*B* reports the fraction of reports submitted by the monitors measured as the number of weeks in which reports were submitted by at least one monitor in the CBWMO divided by the total number of posttreatment weeks. In most treated communities, monitors submitted at least one report weekly. On average, a report was submitted in 76% of the posttreatment weeks. Fig. 2*C* reports the fraction of summary reports read by the

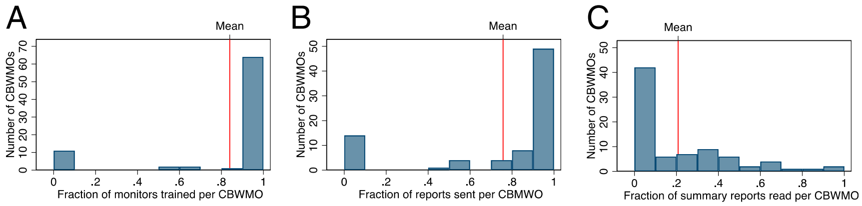

Fig. 2. Distribution of fraction of monitors trained, fraction of reports sent by monitors, and fraction of summary reports read by committee members. The fraction of monitors trained per CBWMO (*A*) is the number of monitors trained divided by the number of monitors assigned to that CBWMO. The fraction of reports sent per CBWMO (*B*) is the ratio of the number of weeks when at least one report was sent by a monitor with respect to the total number of weeks after treatment. The fraction of summary reports read per CBWMO (*C*) is the ratio of the number of weeks in which at least one committee member read the summary report with respect to the total number of weeks after treatment. Histograms are constructed using all treated CBWMOs.

CBWMO management committee measured as the number of weeks in which at least one committee member read the summary report divided by the total number of posttreatment weeks. In contrast to the high compliance among monitors, compliance among management committees was low: On average, a report was read in 21% of the posttreatment weeks.

In the PAP, we defined one binary indicator of compliance for use in estimating average treatment effects on compliant communities; committee members in a CBWMO complied with the treatment if at least one committee member opened a summary report in 75% or more of the weeks. Here, we also add a second binary indicator of compliance that captures whether monitors submitted their reports (i.e., the second process step in our definition of a "working" monitoring system): at least one of the monitors sent a report in 75% or more of the weeks. Among the 69 CBWMOs in which at least one monitor was trained, 58 (84%) had at least one of the monitors send a report in 75% or more of the weeks. However, only 4 (5.8%) of the CBWMOs had a committee member read a report in 75% or more of the weeks.

Estimated Effects of Community Monitoring on the Primary Outcome Variables. The estimated effects of the program on the four primary outcome variables are reported in Fig. 3 (for more details, see *Materials and Methods*). We estimated the intent-to-treat effects (ITT), which are the effects of the treatment "as assigned," ignoring noncompliance. The signs of the estimated effects are consistent with the hypotheses that arise from the theory of change, except in the case of user beliefs about their community wasting water. The estimated impact on monthly electricity consumed in pumping (H1) is a reduction of 7.9% (0.06 SD). The estimated impact on the presence of bacteria (H2) is a reduction of 6.2% (0.15 SD). The estimated impact on the user satisfaction with water service (H3-1) is an increase of 1.4% (0.08 SD). For these three estimated effects, the 95% confidence intervals include zero, but most of the interval values are in the hypothesized direction. The estimated program impact on user beliefs about their community wasting water (H3-2) is an increase of 0.8% (0.005 SD), which is small and much less precisely estimated than the other effects.

Estimated Effects of Community Monitoring in Compliant Communities. Assuming the training workshop did not have a direct effect on the primary outcomes independent of the monitoring, we can use our binary indicators of monitor and committee compliance to estimate the effect of the program in compliant communities. These complier average causal effect (CACE) estimates are reported in Fig. 4. The estimated CACEs are larger than the estimated ITT effects, albeit even less precisely estimated (perhaps unsurprising given that the complier groups are smaller, particularly for CBWMOs with committees that regularly read the summary reports). The randomized treatment assignment is a weak instrument for inducing committee members to read most of the submitted reports, and thus the estimator for this CACE may be substantially biased.

Effect of the Program on Intermediate Variables. To evaluate the mechanisms through which the program may have affected outcomes, we estimate the effect of the program on twelve intermediate variables. Table 2 shows the signs of the ITT effects in comparison with the theory of change (*SI Appendix*).

For 9 out of the 12 intermediate variables, the estimated effects are in the predicted direction. The signs of the estimates are consistent with the theory of change predictions that: 1) the program would increase the committee members' and the community's perception that they are under scrutiny; 2) the program would increase community interest about the water system and its management (more suggestions and complaints about the system); and 3) the program would improve CBWMO financial management and maintenance expenditures on infrastructure and water quality tests as well as reduce the number of overflowing water tanks (a form of leaks).

Only for the last intermediate outcome (overflowing tanks) can the null hypothesis of zero effect be rejected ($P < 0.05$; unadjusted for multiple comparisons). The estimated impact on the binary variable "CBWMO has at least one tank that overflows" is a reduction of 73% reporting "Yes" (0.32 SD).

For three intermediate variables, the estimated effects are not in the predicted direction. The signs of the estimated effects are not consistent with the theory of change predictions that: 1) the program would increase attendance at the CBWMO public assemblies[§]; 2) the program would increase the number of CBWMOs with a plumber working at least 50% of the time; and 3) the program would increase the number of conversations about water problems.

Only for this last outcome measure can the null hypothesis of zero effect be rejected ($P < 0.05$; unadjusted for multiple comparisons). We estimate a 47% (0.12 SD) decrease in the number of conversations about water.

To shed further light on the results in Table 2, we also report the values of two of the process variables measured in the evaluation: the proportion of households who reported being aware of the monitoring system and the proportion who reported accessing the summary reports. On average, only 12% of the households reported knowing about the monitoring system and only 2% reported accessing the summary reports. Thus, the channels in the theory of change that rely on broad community engagement with water issues are unlikely to have been widely operative. Despite the CBWMO committees' communication efforts and the public workshops, knowledge about the monitoring system did not spread as expected by the project team.

[§] We have two sources of data on attendance at CBWMO public assemblies: the survey of committee members and the survey of households. Using data from the committee surveys, the estimated effect is in the anticipated direction. Using data from household surveys, the estimated effect is in the opposite direction. Given these contradictory results, our interpretation is that the estimated effects are not consistent with the theory of change predictions.

Table 2. Evaluating program effects along the causal pathway

Mechanism	Variable	Expected effect	Results Coef. sign	 Rejected H0: b = 0 H1: b ≠ 0 ($P \leq 0.05$)
Committees perceive greater scrutiny of their performance by users	Committee's beliefs about community paying attention	Positive	✓	✗
Community has more interest in water system and committee performance	Attendance at public assemblies (committee members' survey)	Positive	✓	✗
	Attendance at public assemblies (household survey)	Positive	✗	✗
	Number of suggestions by community members	Positive	✓	✗
	Community members' conversations about water issues	Positive	✗	✓
Community perceives greater scrutiny of illegal or improper water use	Community members' beliefs about punishment probability for an illegal water connection	Positive	✓	✗
Committees improve their water pricing	Marginal price of water	Positive	✗	✗
Committees improve their water system management	Index of management practices	Positive	✓	✗
Committees improve infrastructure maintenance	Plumber works more than 50% of the time	Positive	✗	✗
	Maintenance expenditure	Positive	✓	✗
	Number of water quality tests	Positive	✓	✗
	At least one water tank leaks/overflows	Negative	✓	✓

In column 4, the coefficient sign shows a check (✓) when it coincides with the expected effect from column 3.

Cost-Effectiveness Analysis. We estimate the cost-effectiveness of the water monitoring system and compare it with the cost-effectiveness of a different nonmonetary strategy to address water scarcity: encouraging the adoption of water-efficient technology (WET) (40). Two years prior to the launch of the monitoring system, CATIE ran a WET program in nine rural communities located in the same region. The program was implemented as a randomized controlled trial in which interested households were block randomized into either a group that received free WET or a group that stayed with the status quo technologies (four of these communities were also treated CBWMOs in the monitoring experiment; two were part of the control group, and three were not part of the experiment). The WET installation comprised a package of efficient showerheads (1.5 gallons per minute) and faucet aerators (1 gallon per minute) for bathroom and kitchen faucets.

We compare the cost-effectiveness of both programs over a 5 y period.[¶] The average program cost of installing the water-efficient technologies in each house was US$33.6, and the technologies reduced monthly water consumption by 9.1% (40). The estimated average discounted cost of establishing and maintaining the monitoring system during a 5 y period is $3,480 per CBWMO. The main assumption needed to calculate the reduction in water consumption due to the monitoring program is that the program's estimated percentage reduction in electricity consumed in pumping corresponds to the percentage reduction in groundwater pumping. We believe our assumption is plausible because the monitoring program did not induce changes in the energy efficiency of the pumps, nor would the modest estimated treatment effect likely change the marginal cost of pumping a cubic meter (m^3) of groundwater. Other assumptions used in these calculations are described in *SI Appendix*.

We calculate that for every dollar (USD) invested, the WET program reduced groundwater pumping by 7.9 m^3, and the monitoring program reduced groundwater pumping by 14.0 m^3 over a 5 y period. The programs would be equally cost-effective only if water reduction in the monitoring program falls to 56% of the estimated effect in electricity for pumping (7.9%) (*SI Appendix*). However, we note that the effect of the WET program is more precisely estimated than the effect of the monitoring program. While the confidence interval of the estimated reduction per dollar invested in the WET program is between 5 m^3 and 10.8 m^3, the confidence interval of the estimated effect per dollar invested in the monitoring program is between a 32.7 m^3 reduction and a 4.8 m^3 increase. Thus, our cost-effectiveness analysis should be considered suggestive rather than definitive.

Discussion

The literature on common pool resources has emphasized the need for studies that provide causal analysis of the effects of the conditions (1) that presumably enable sustainable common pool resource management (11). However, when these conditions arise endogenously, researchers face challenges in disentangling the effect of one condition separately from other community attributes that are correlated with both the condition and the use of the common pool resource. Even if causal inference in such contexts were straightforward, the inferences drawn may not generalize to communities where the enabling conditions do not exist and outside actors are seeking to encourage their formation. Externally driven monitoring programs have been studied in the literature on community-based environmental monitoring, but this literature has not provided empirical evidence of the causal effects of monitoring on environmental conditions. Our study contributes to both literatures by exploring the effects of an externally driven community-based monitoring program on water quantity, water quality, and user satisfaction and by exploring the mechanisms through which the program affects these outcomes.

The estimated impacts of the program 1 y after its start were modest (~0.10 SD) and imprecisely estimated, but they were in the desired directions: Communities with monitors pumped less groundwater from the aquifer, had better water quality, and had higher customer satisfaction. Using the estimated effect on groundwater pumping, we find that the cost-effectiveness of community monitoring compares favorably to an alternative, demand-based intervention that relies on household water-efficient technology adoption to reduce groundwater pumping. Although the cost-effectiveness

---

[¶] This is the average useful lifespan of the WET estimated by the WET program participants (40).

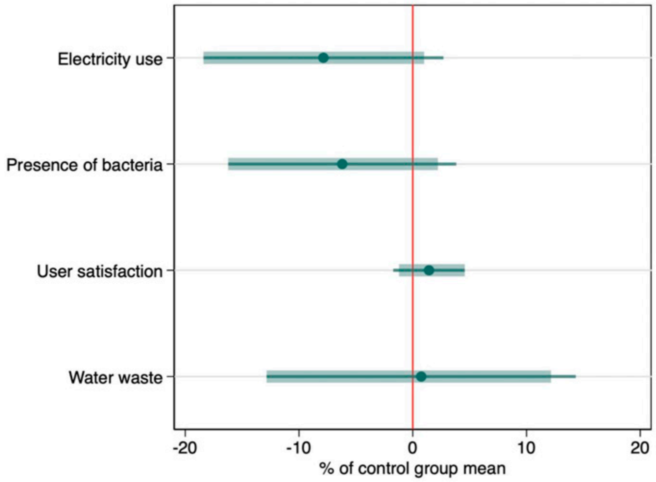

Fig. 3. Estimated effects (ITTs) of the program on the primary outcome variables. The point indicates the effect as a percentage of the control group mean, and the thin segments represent the 95% confidence intervals for two-tailed hypotheses. The ends of the thick segments that fully cover the thin segments indicate the direction of the prespecified one-tailed hypotheses (the other end represents the one-sided 95% CI bound). If the thick segments do not bound zero, we reject null hypotheses at the $\alpha = 0.05$ level.

results are sensitive to assumptions, they imply that more testing of community monitoring is warranted.

Our theory of change posited that the program's effects on the outcome variables would be mediated by an increase in the management committees' and citizens' perception that they are under scrutiny, by an increase in community interest about the water system's management, and by an increase in new information with which the CBWMO committees could improve services. Using intermediate outcome variables, we find evidence consistent with these conjectures, but the effects of the program on these variables are small and imprecisely estimated.

As noted by others (37), creating and sustaining a monitoring program is a "second-order" collective action problem: The benefits of the monitoring program are public, while the costs of participating in it are private. The supply of monitoring will thus be suboptimal in most cases, and it may be worse in cases in which the original motivation for creating the monitoring program comes from external actors.

We see evidence of this collective action challenge in our study context. The program was more successful in engaging monitors than engaging the management committee members. In the literature on community-based environmental monitoring, low use of the monitoring data by managers or policy makers is often attributed to a lack of trust in the data (41). However, when we asked committee members why they did not read the summary reports, only 6% of respondents listed "lack of trust in data." Instead, around one-third reported not having time to read the summary reports or simply forgetting that the reports were available. We posit that committee engagement could be enhanced through two channels. First, committees may have been more likely to read the summary reports if they had played a role in designing them (42). Second, committees may have been more likely to read the reports if their community members held them more accountable for their performance. Greater accountability could be achieved by integrating presentations of the summary reports by monitors at management committee meetings or public assemblies, during which committee members would be expected to explain their strategies to improve the water system. To further provide motivations for accountability, the monitor reports could also include comparisons of monitoring indicators across time or across communities. In comparison to the committee members, resource users in the community were even less likely to read the summary reports or even know the program existed. We believe that monitoring programs should do more community advertising of the program to ensure users are aware of its presence and purpose.

Although monitors were more engaged than committee members and users, sustaining their engagement was a challenge. The total number of weekly reports decreased over time (Fig. 5) despite the payments for each report, the automatic weekly reminders, and personal phone contact by the project team after several weeks had passed without a report. Although we cannot determine whether the declines in weekly reporting reduced the overall estimated impact of the monitoring program, a program cannot be effective if monitors cease monitoring. Our study demonstrates that a fixed payment is not enough to sustain monitor motivation. To keep monitors engaged, a program may

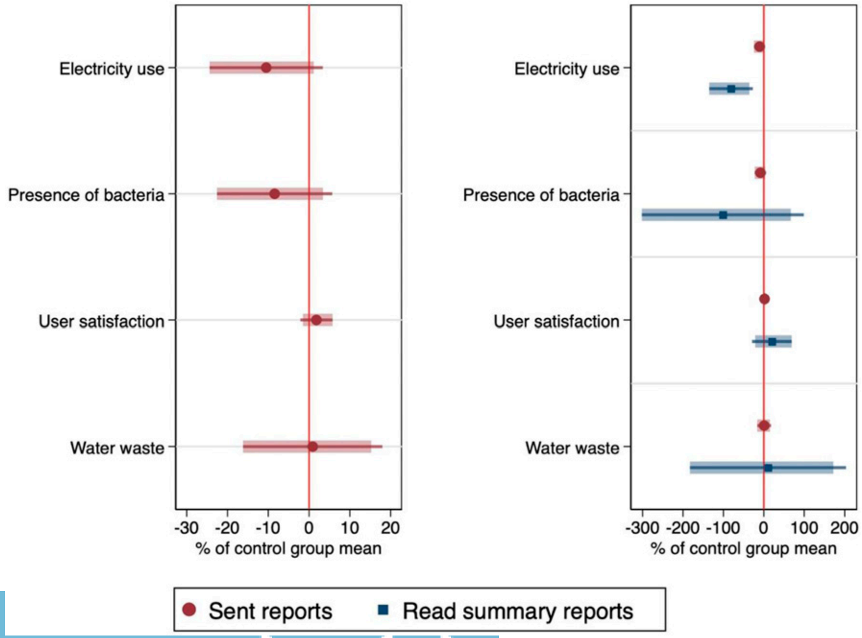

Fig. 4. Effects on the primary outcome variables among two types of compliers. The graph on the left corresponds to the "Sent reports" compliers. The one on the right compares these estimates with the ones of the "Read summary reports" compliers. The point indicates the effects as a percentage of the control group mean, and the thin segments represent the 95% confidence intervals of two-tailed hypotheses. The ends of the thick segments that fully cover the thin segments indicate the direction of the prespecified one-tailed hypotheses (the other end represents the one-sided 95% CI bound). If the thick segments do not bound zero, we reject null hypotheses at the $\alpha = 0.05$ level.

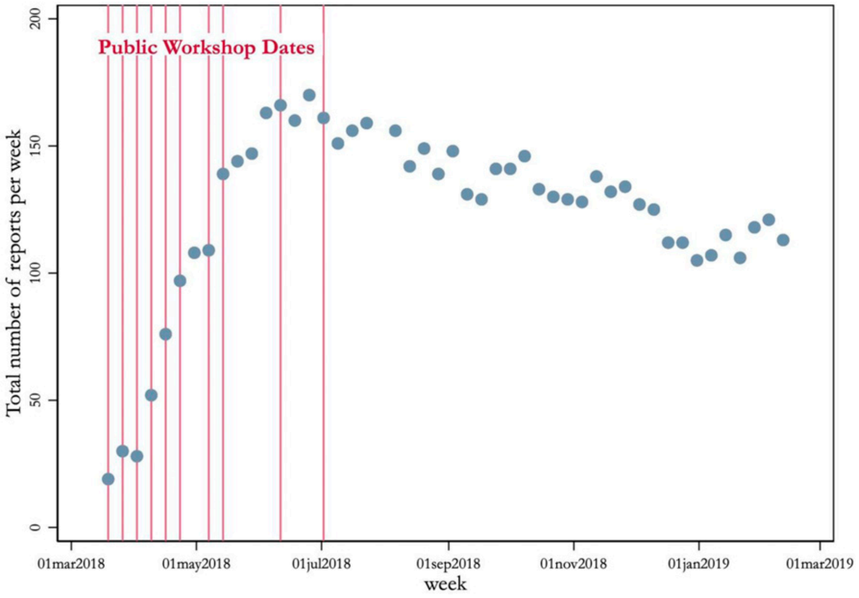

Fig. 5. Total number of weekly monitor reports. The figure shows the total number of reports sent by monitors from all the CBWMOs each week. The red vertical lines indicate the weeks when public workshops took place: from the third week of March through end of June.

wish to try providing the following: 1) more exciting incentives that vary over time and are based on a mix of luck and performance; 2) more frequent positive feedback, encouragement, and recognition for monitors; and 3) a stronger sense of a common purpose and civic duty among the monitors.

One potential path for overcoming the challenges described in the previous paragraphs is to increase community engagement in the design and establishment of the community monitoring program. If such engagement were effective in overcoming collective action problems in community monitoring, then our results could be viewed as underestimates of the impacts of community monitoring when it is codesigned or arises endogenously.

In addition to the collective action incentive problem that all monitoring programs face, a technology-based monitoring program like the SIMA app faces additional challenges. Although technology can lower the costs of collecting and disseminating monitoring data, it can also raise the costs of broadening community participation in the program. A community-based monitoring based on a smartphone app requires that users own a smartphone, that they have access to Internet, and that they know how to use their smartphones and the app. Although all monitors had a cell phone and knew how to use it properly, only 61% of committee members in the public workshop had a smartphone and, of those with a smartphone, only 87% of them could download the SIMA app. Moreover, many committee members had problems using their smartphones, and the great majority of them did not reach out to the project team for help. Forty-two percent of them indicated technology problems—they could not install the app, lost the password, or did not know how to use apps on their phone. It is possible that resource users had the same problems. Had the program anticipated this problem, a potential solution would have been to hold training sessions for community members to teach them to read the reports via the app. The project team could have also called committee members periodically to check whether they had any problems using the app and to offer help rather than waiting until members contacted the project team. A less expensive option may have been for the project team to print, or encourage a committee member to print, the summary reports, so that the information could be shared with all the members of the CBWMO committee and with the rest of the community.

In summary, the experimental results provide some support for claims that an externally driven, smartphone-based community-based monitoring program can cost-effectively improve common pool resource governance. Yet we also detect collective action challenges and technology barriers to engaging the relevant actors and, given the short postintervention period of our study, we cannot assess how sustainable the program's impacts may be. Replications of our design can help increase the precision of the impact estimates, assess the impacts of variations in program attributes and delivery, and assess the generalizability of the impacts to other contexts and longer time periods.

Materials and Methods

The PAP, the data, and code to reproduce the results in this study are available at https://osf.io/bmndv/. Ethical approval was obtained from Johns Hopkins University (HIRB00008760). All participants in the household survey provided informed consent.

Statistical Models. Table 3 presents all the models used for the ITT estimations of the main outcome variables and intermediate variables. To estimate the ITT on electricity consumed in pumping, we use a random effects model:

$$Y_{it} = \alpha + \beta T_{it} + \theta X_i + \varepsilon_i + \epsilon_t + \mu_{it}, \qquad [1]$$

where Y_{it} is the logarithm of electricity consumed in pumping of CBWMO i in month t, T_{it} is 1 if CBWMO i is assigned to treatment, and t is a post-treatment month, X_i are the CBWMO-level control variables, ε_i is the CBWMO-specific effect that, given randomized treatment assignment, is assumed to be a random variable and uncorrelated with the treatment, ϵ_t is the month-specific effect, and μ_{it} is the idiosyncratic error term. Heteroskedastic-robust SEs are clustered at the CBWMO level. The CBWMO-level control variables are the block randomization variables (*SI Appendix*) and the variable severity of the water problem.

To estimate the ITT on the presence of bacteria, we use a probit model:

$$Prob(Y_{ki}) = \phi(\alpha + \beta t_i + \theta X_i + \mu_{ki}), \qquad [2]$$

where Y_{ki} is a binary variable that indicates whether fecal coliforms were found in the water quality test k in CBWMO i, t_i is 1 if CBWMO i is assigned to treatment, X_i is the same as in Eq. **1**, and μ_{ki} is the error term. A water quality test was conducted at three points in the water infrastructure in each community. Heteroskedastic-robust SEs are clustered at the CBWMO level.

To estimate the ITT on the user satisfaction index, we use an Ordinary Least Squares model:

Table 3. Statistical models for ITT estimations

Variable	Model	Unit of analysis	Pretreatment values
Main outcome variables			
Log of monthly electricity consumed in pumping	Random Effects	CBWMO month	Yes
Presence of bacteria	Probit	CBWMO water test	No
User satisfaction with water service	OLS	Household	No
User belief about community wasting water	Probit	Household	No
Intermediate variables			
CBWMO's beliefs about community paying attention	Ordinal Probit	CBWMO	Yes
Attendance to public assemblies (committee's report)	Negative Binomial	CBWMO	Yes
Attendance to public assemblies (community's report)	Probit	Household	No
Number of community's suggestions	Negative Binomial	CBWMO	Yes
Community's conversations about water issues	Negative Binomial	Household	No
Community's beliefs about punishment probability for an illegal water connection	Ordinal Probit	Household	No
Marginal price of water	OLS	CBWMO	Yes
Index of management practices	Fractional Response GLM	CBWMO	Yes
Plumber works more than 50% of the time	Probit	CBWMO	Yes
Maintenance expenditure	OLS	CBWMO	Yes
Number of water quality tests	Poisson	CBWMO	Yes
At least one water tank overflows	Probit	CBWMO	Yes

$$Y_{ji} = \alpha + \beta t_i + \theta X_i + \gamma Z_{ji} + \mu_{ji}, \qquad [3]$$

where Y_{ji} is the user satisfaction index of household j in CBWMO i, Z_{ji} are the household-level control variables, and μ_{ji} is the error term. The household-level control variables are sex and age of responder, the number of family members, the number of years in the community, and dummy variables for religion, education level, and type of house ownership. Heteroskedastic-robust SEs are clustered at the CBWMO level.

To estimate the ITT on the users' belief about their community wasting water, we use a probit model:

$$Prob(Y_{ji}) = \phi\left(\alpha + \beta t_i + \theta X_i + \gamma Z_{ji} + \mu_{ji}\right), \qquad [4]$$

where Y_{it} is a binary variable that indicates whether household j in CBWMO i believes that the community wastes water, and the other attributes of the model are as in Eq. **3**.

In the case of the intermediate variables, some of the estimations are at the CBWMO level, and some are at the household level. All CBWMO-level estimations include the corresponding pretreatment outcome variable and the CBWMO-level control variables, and we apply heteroskedastic-robust SEs. All household-level estimations include control variables at both the CBWMO level and at the household level, and we clustered the heteroskedastic-robust SEs at the CBWMO level (we do not have pretreatment outcome variables).

We estimate CACEs for all the main outcome variables. In the case of the monthly electricity consumed in pumping, we apply a two-stage least-squares random effects model and use the treatment assignment variable in the posttreatment period as instrument. For the other main outcome variables—presence of bacteria, user satisfaction with water service, and user satisfaction with community water usage—we use a two-stage least-squares model and the treatment assignment variable as instrument.

Table 4 shows the ITT and CACE results for all the main variables. We report the raw estimates or the marginal effects, the *P* values, the number of observations, and the estimate in percentages with respect to the control group mean, the corresponding *P* value, and the estimate as a proportion of the control group SD. We also include a test for a weak instrument. In the case of the monthly electricity consumed in pumping, we report the F-statistic of excluded instruments. As a rule of thumb, a value below 10 indicates a weak instrument. For the other variables, we apply the weak instrument test of Montiel Olea and Pflueger (43). This test rejects the null hypothesis of weak instrument when the effective F-statistic exceeds a critical value, which depends on the significance level and the chosen weak instrument threshold value, τ. In all the models, we cannot reject the weak instrument hypothesis when the instrumented

Table 4. ITT and CACE estimates for main outcome variables

								Weak instrument test	
Main outcome variable	Model	Estimate	*P* value	N	Percentage	*P* value	Number of SDs	F/effective F	20% critical value
Log of monthly electricity consumed in pumping	ITT	−0.08	0.16	7420	−7.85%	0.14	−0.06		
	Sent reports	−0.11	0.16	7420	−10.51%	0.14	−0.08	217.71	
	Read reports	−1.66	0.25	7420	−81.06%	0.00	−0.59	4.13	
Presence of bacteria	ITT	−0.05	0.23	417	−6.20%	0.23	−0.15		
	Sent reports	−0.07	0.24	423	−8.45%	0.24	−0.21	262.26	15.06
	Read reports	−0.87	0.32	423	−101.36%	0.32	−2.52	4.63	15.06
User satisfaction with water service	ITT	0.06	0.37	3799	1.43%	0.37	0.08		
	Sent reports	0.07	0.37	3799	1.80%	0.37	0.09	297.28	15.06
	Read reports	0.77	0.43	3799	19.87%	0.43	1.05	5.16	15.06
User belief about community wasting water	ITT	0.00	0.91	3551	0.75%	0.91	0.00		
	Sent reports	0.00	0.92	3551	0.90%	0.92	0.01	306.59	15.06
	Read reports	0.03	0.92	3551	10.13%	0.92	0.07	5.07	15.06

Columns 3 to 5 show the raw estimates or the marginal effects, *P* values for two-tailed hypotheses, and number of observations. Columns 6 and 7 show the estimates in percentages of the control group mean and the *P* values for two-tailed hypotheses. In column 8, we report the estimates in number of SDs of the control group. In columns 9 and 10, we report the results of testing whether the excluded instruments are weak. For the logarithm of monthly electricity consumed in pumping, we report the F-statistic of excluded instruments. For the rest of the other variables, we report the Montiel–Pflueger effective F-statistic and the critical value. If the effective F-statistic is below the critical value, the instrument is weak.

variable is the CBWMO committee members' compliance with reading most of the summary reports; however, we can reject it when the instrumented variable is monitors' compliance with sending most of the reports.

In our PAP, we also specified a secondary outcome measure: a water quality index constructed from six water quality variables (turbidity, chlorine, alkalinity, pH, oxygen, and temperature). We construct the index using Eq. **2**. The estimated impact on the secondary water quality variable is an increase of 15.7% (0.32 SD), which is the opposite direction of the hypothesis, and the 95% confidence interval does not bound zero. The positive effect is driven by the alkalinity value, which is a persistent feature of the water source and thus should not have been included in the definition of the index in the PAP. Once we remove this variable, the estimated effect reduces to 4.7% (0.10) and is not statistically significant.

Deviations from Registered PAP. We prespecified an analysis of moderator effects (i.e., subgroup effects or conditional intent-to-treat effects). Because our estimated intent-to-treat effects were imprecisely estimated, we do not believe estimating conditional intent-to-treat effects is prudent, and thus we abandon that planned analysis.

We prespecified using a Poisson model for all count data variables, but after obtaining the data, a test of goodness of fit implied that the Negative Binomial model was a better fit for some of the intermediate outcome variables. In those cases, we used the Negative Binomial model as described in Table 3.

In the PAP, we failed to include text about the natural logarithm transformation of the electricity data, which was applied because of the highly skewed distribution of that variable.

In the PAP, we failed to include text referring to a control variable used in the household-level estimators: gender of respondent.

We prespecified the second process variable (Fig. 2) as the fraction of reports sent with respect to the total number of reports that could have been sent. Since a working monitoring system requires that at least one report is submitted by the monitors every week, we changed the process variable to the number of weeks in which reports were submitted by at least one monitor in the CBWMO divided by the total number of posttreatment weeks. This new variable is also more comparable to the third process variable (fraction of summary reports read by the CBWMO management committee).

We added a second binary indicator of compliance that captures whether monitors submitted their reports.

Data Availability. Anonymized dta files are available via the Open Science Framework and can be accessed at https://osf.io/bmndv/ (38).

ACKNOWLEDGMENTS. We thank the UK's Department for International Development and the Evidence in Governance and Politics network for funding the study and the intervention. We thank the field team led by Maureen Arguedas and Adrian Vargas from CATIE for implementing the study.

1. E. Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press, 1990).
2. F. Danielsen *et al*., Community monitoring for REDD+: International promises and field realities. *Ecol. Soc.* **18**, 41 (2013).
3. IPCC, "Climate change 2014: Mitigation of climate change" in *Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, O. Edenhofer *et al*., Eds. (Cambridge University Press, Cambridge, United Kingdom, 2014).
4. C. C. Conrad, K. G. Hilchey, A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environ. Monit. Assess.* **176**, 273–291 (2011).
5. G. Whitelaw, H. Vaughan, B. Craig, D. Atkinson, Establishing the Canadian community monitoring network. *Environ. Monit. Assess.* **88**, 409–418 (2003).
6. M. Björkman, J. Svensson, Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda Community-Based Monitoring In Uganda* *Q. J. Econ.* **124**, 735–769 (2009).
7. E. A. Coleman, Institutional factors affecting biophysical outcomes in forest management. *J. Policy Anal. Manage.* **28**, 122–146 (2009).
8. C. C. Gibson, J. T. Williams, E. Ostrom, Local enforcement and better forests. *World Dev.* **33**, 273–284 (2005).
9. E. Ostrom, H. Nagendra, Insights on linking forests, trees, and people from the air, on the ground, and in the laboratory. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19224–19231 (2006).
10. S. L. Morgan, C. Winship, *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (Cambridge University Press, 2014).
11. A. Agrawal, Common property institutions and sustainable governance of resources. *World Dev.* **29**, 1649–1672 (2001).
12. E. Ostrom, The value-added of laboratory experiments for the study of institutions and common-pool resources. *J. Econ. Behav. Organ.* **61**, 149–163 (2006).
13. J. Henrich *et al*., Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
14. E. Fehr, S. Gächter, Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).
15. G. Shreedhar, A. Tavoni, C. Marchiori, Monitoring and punishment networks in an experimental common pool resource dilemma. *Environ. Dev. Econ.* **25**, 66–94 (2019).
16. B. L. Sullivan *et al*., Using open access observational data for conservation action: A case study for birds. *Biol. Conserv.* **208**, 5–14 (2017).
17. D. C. McKinley *et al*., Citizen science can improve conservation science, natural resource management, and environmental protection. *Biol. Conserv.* **208**, 15–28 (2017).
18. F. Danielsen, N. D. Burgess, M. Jensen, K. Pirhofer-Walzl, Environmental monitoring: The scale and speed of implementation varies according to the degree of people's involvement. *Appl. Ecol.* **47**, 1166–1168 (2010).
19. M. Gharesifard, U. Wehn, P. van der Zaag, What influences the establishment and functioning of community-based monitoring initiatives of water and environment? A conceptual framework. *J. Hydrol. (Amst.)* **579**, 124033 (2019).
20. B. A. Olken, Monitoring corruption: Evidence from a field experiment in Indonesia. *J. Polit. Econ.* **115**, 200–249 (2007).
21. T. Andrabi, J. Das, A. I. Khwaja, Report cards: The impact of providing school and child test scores on educational markets. *Am. Econ. Rev.* **107**, 1535–1563 (2017).
22. K. Donato, A. Garcia Mosqueira, Information improves provider behaviour: A replication study of a community-based monitoring programme in Uganda. *J. Dev. Stud.* **55**, 967–988 (2019).
23. E. Duflo, P. Dupas, M. Kremer, School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools. *J. Public Econ.* **123**, 92–110 (2015).
24. M. B. Nyqvist, D. de Walque, J. Svensson, Experimental evidence on the long-run impact of community-based monitoring. *Am. Econ. J. Appl. Econ.* **9**, 33–69 (2017).
25. M. M. Mekonnen, A. Y. Hoekstra, Sustainability: Four billion people facing severe water scarcity. *Sci. Adv.* **2**, e1500323 (2016).
26. S. M. Gorelick, C. Zheng, Global change and the groundwater management challenge. *Water Resour. Res.* **51**, 3031–3051 (2015).
27. H. Vaux, Groundwater under stress: The importance of management. *Environ. Earth Sci.* **62**, 19–23 (2011).
28. Fundación Avina CLOCSAS, *La Asociatividad entre Organizaciones Comunitarias de Servicios de Agua y Saneamiento -OCSAS- en Latinoamérica* (Fundación Avina, CLOCSAS, Panama, 2017).
29. H. Lockwood, G. Mansour, S. Smits, S. Smets, *Global Study on Sustainable Service Delivery Models for Rural Water: Evidence from 16 Countries Introduction and Methodology* (Loughborough University, 2017).
30. R. Madrigal-Ballestero, M. A. Naranjo, Adaptive capacity, drought and the performance of community-based drinking water organizations in Costa Rica. *J. Water Clim. Chang.* **6**, 831–847 (2015).
31. L. Sánchez Hernández, *Situación Actual de las Organizaciones Comunales Prestadoras de Servicios de Agua Potable y Saneamiento* (Programa Estado de la Nación, Costa Rica, 2019).
32. P. Imbach, *et al*., "Impacts of climate change on ecosystem hydrological services of Central America: Water availability" in *Climate Change Impacts on Tropical Forests in Central America. An Ecosystem Service Perspective*, A. Chiabai, Ed. (Routledge, 2015), pp. 65–90.
33. Contraloría General de la República, "Informe sobre la auditoria operativa acerca de la eficacia y eficiencia del Instituto Costarricense de Acueductos y Alcantarillados para garantizar la prestación del servicio de abastecimiento de agua potable" (Informe Nro. DFOE-AE-IF-08-2012, Contraloría General de la República, Costa Rica, 2012).
34. L. Ávila, *El agua que se pierde* (EL Financ, 2019).
35. R. Madrigal, F. Alpízar, A. Schlüter, Determinants of performance of community-based drinking water organizations. *World Dev.* **39**, 1663–1675 (2011).
36. M. Golden, D. Green, D. L. Nielson, D. Rubenson, T. Slough, EGAP Metaketa III: Natural Resource Governance. Pre-Analysis Plan (2020). https://osf.io/5pvud. Accessed 11 May 2021.
37. A. Agrawal, S. Goyal, Group size and collective action: Third-party monitoring in common-pool resources. *Comp. Polit. Stud.* **34**, 63–93 (2001).
38. M. Bernedo Del Carpio, F. Alpízar, P. J. Ferraro, Community-based monitoring to facilitate water management by local institutions in Costa Rica. Open Science Framework. https://osf.io/bmndv/. Deposited 4 May 2021.
39. F. Alpízar, M. Bernedo Del Carpio, P. J. Ferraro, B. S. Meiselman, The impacts of a capacity-building workshop in a randomized adaptation project. *Nat. Clim. Chang.* **9**, 587–591 (2019).
40. F. Alpizar, M. Bernedo Del Carpio, P. J. Ferraro, *Input Efficiency as a Solution to Externalities: A Randomized Controlled Trial* (Environment for Development, 2021).
41. U. Wehn, A. Almomani, Incentives and barriers for participation in community-based environmental monitoring and information systems: A critical analysis and integration of the literature. *Environ. Sci. Policy* **101**, 341–357 (2019).
42. A. J. McKay, C. J. Johnson, Identifying effective and sustainable measures for community-based environmental monitoring. *Environ. Manage.* **60**, 484–495 (2017).
43. J. L. M. Olea, C. Pflueger, A robust test for weak instruments. *J. Bus. Econ. Stat.* **31**, 358–369 (2013).
44. E. Estrada Hernandez, A. Chinchilla Medina, *Acceso y Uso de los Servicios de Telecomunicaciones en Costa Rica 2017* (Ministerio de Ciencias, Tecnología y Comunicaciones, Costa Rica, 2019).
45. MTSS, Decreto Nr. 40743-MTSS (Ministerio de Trabajo y Seguridad Social, 2017).