Community Participation in Decision-Making
Evidence from an experiment in providing safe drinking water in Bangladesh

Malgosia Madajewicz*[1], Anna Tompsett[2] and Ahasan Habib[3]

[1]NASA Goddard Institute for Space Studies, Columbia University, New York
[2]School of International and Public Affairs, Columbia University, New York
[3]NGO Forum for Public Health, Dhaka, Bangladesh

October 3, 2014

Abstract

The hypothesis that participation in decision-making by intended beneficiaries of social programs improves the outcomes of those programs has long been influential in the academic literature and in policy. This paper presents the first experimental evidence on the effect of transferring decision-making authority to targeted beneficiaries on the impact of a project to provide a local public good. We randomly assigned participatory and non-participatory decision-making structures to communities who received an otherwise identical intervention, a package of technical advices and subsidies to provide safe drinking water sources. Participation in decision-making resulted in larger reported increases in access to safe drinking water, but only when we imposed rules on the decision-making process that were designed to limit the appropriation of project benefits by elite or influential groups or individuals. Villages in which communities participated in decision-making under rules designed to prevent appropriation reported a significantly greater increase in access to safe drinking water (an increase of 25%) relative to villages in which project staff took decisions (14%). In villages in which the communities participated in decision-making without imposed rules, the change in access to safe drinking water was the same (14%) as in villages in which project staff took decisions. We conclude that participation can improve impact in economically important respects; that the risk of appropriation in this context is real and important, and that the rules we applied to limit appropriation – minimum representation requirements and decision by unanimous consensus – were effective in accomplishing their objective.

*E-mail: *mm1174@columbia.edu*. Funded by the National Science Foundation Grant 0624256 *DHB: Decentralization and Local Public Goods: How Does Allocation of Decision-making Authority Affect Provision?*. Thanks to all at NGO Forum for Public Health for their kindness and hospitality as well as their collaborative efforts, in particular Executive Director Mr S.M.A. Rashid. Kabid Ahmed, Md. Shamsul Karim, Fatema Tuz-Zohora, and the indomitable Md. Rofi Uddin (Robi) provided excellent research assistance. Thanks to seminar participants at Columbia University, Stockholm University and the CPER—PODER Development Symposium for helpful questions and comments.

1 Introduction

The hypothesis that participation in decision-making by intended beneficiaries of social programs improves the outcomes of those programs has been influential in the academic literature and in policy for some time (e.g. Stiglitz, 2002; World Bank, 2004). Advocates of the policy argue that involving communities in project decision-making has multiple benefits: improving project targeting, by drawing on information available to the community but not to outsiders; increasing 'buy–in' and generating a 'sense of ownership' of the project, thereby improving long-term management and increasing maintenance of program assets; and promoting transparency and accountability in project delivery. However, participation is costly, both to beneficiary communities and to donor organizations, and programs in which communities participate in decision-making may be more susceptible to the 'capture' of project benefits by elite or influential community members[1].

Much of the early evidence in support of this hypothesis was based on cross-sectional analyses[2], case studies[3], or was simply anecdotal. Since the choice of a decision-making structure is likely to be otherwise correlated with project, community and implementing agency characteristics, identification of causal effects is difficult and sensitive to critical assumptions. This paper presents the first experimental evidence on whether transferring decision-making authority to intended beneficiaries affects the impact of a project to provide a local public good.

We randomly assigned different decision-making structures to communities who received an otherwise identical intervention, a package of subsidies and technical advice to provide up to three sources of safe drinking water. Many rural Bangladeshi communities currently use sources of water that are susceptible to arsenic or, less commonly, bacterial contamination. The fixed cost of installing arsenic-safe drinking water sources is high relative to local incomes, while the marginal cost of producing more safe water is low. As a result, provision of safe drinking water presents a common pool resource or local public good problem. The random assignment ensured that the communities in which we implemented the project under different decision-making structures were

[1]See Mansuri and Rao (2013) for a comprehensive review.

[2]Examples include: Isham, Narayan, and Pritchett (1995), Sara and Katz (1997), A. Khwaja (2004), Fritzen (2007), A. I. Khwaja (2009).

[3]Examples include: Kleemeier (2000), Fung and Wright (2003), Rao and Ibáñez (2005).

comparable in terms of all other characteristics, allowing us to draw causal inferences about the impacts of the decision-making structures on project outcomes.

The three decision-making structures assigned included a non-participatory decision-making structure and two participatory decision-making structures. In the non-participatory decision-making structure, project staff took all decisions, based on information provided by the community. This process was designed to approximate the way organizations traditionally implemented development projects, where external agents with technical expertise took ostensibly objective decisions. In the first participatory decision-making structure, the community took all decisions using their own internal decision-making processes. This process was designed to approximate the way in which 'participation' is implemented by organizations which place a high value on minimizing interference with local institutions. In the second, we imposed rules on the decision-making process. Under these rules, the community took all decisions by unanimous consensus at a meeting organized by project staff, with requirements imposed for representation of women and the poor. This process was designed to approximate the way in which other organizations implement 'participation', with the active aim of broadening participation in decision-making and reducing marginalization of less-influential groups within the community.

Under all decision-making models, we retained an important participatory component. After decisions were taken, all treatment villages were required to make a contribution towards the cost of water source installation. This cost was fixed beforehand, and was equivalent to 10 and 20% of the total expected cost of water source installation. The communities then had to decide collectively whether or not they would contribute, and how this contribution would be raised. We therefore identify the effects of participation in decision-making over and above the effects generated by any financial contribution.

Overall, the intervention led to an increase in reported access to safe drinking water of 16% relative to a control group. The average treatment effect rises to 18%, compared to a matched control group, when we exclude a subset of villages in which the only feasible technology for providing arsenic-safe drinking water year-round was an arsenic iron removal plant (AIRP). This technology has experienced issues with reliability and effectiveness in the past (Hossain et al., 2005) and our experience suggests that communities strongly prefer tubewells to AIRPs. The treatment effect in the villages in which AIRPs are the only technically feasible option is not statistically different from zero, again compared to a matched control group.

The increase in access to safe drinking water was higher in villages in which the community took decisions and in which decision-making rules were imposed (22% in all villages; 25% if we exclude the AIRP villages) compared to the villages in which project staff took decisions (13%; 14% if we exclude the AIRP villages). However, no differences were observed between the increases in access to safe drinking water when the community took decisions without the imposition of decision-making rules (14%; 15% if we exclude the AIRP villages), and when project staff took decisions. The difference between the change in reported access to safe drinking water in villages in which the community took decisions under imposed rules and the remainder of the treated villages is significant when we remove the villages in which AIRPs were the only option from the analysis. Since the treatment effect is zero in these villages regardless of the structure under which decisions were taken, including them in the analysis is not informative with regards to a comparison between decision-making structures.

We installed an average of 2.1 arsenic safe water sources in each of 127 treatment villages. We installed a slightly larger number of wells in villages in which the community was involved in decision-making (2.2 across both participatory decision-making structures) compared to those in which project staff took decisions (2.0). However, the differences are not statistically significant. Under the non-participatory structure, project staff were instructed to propose locations for water sources in public spaces wherever feasible in order to facilitate access to the sources. Under the participatory structures, communities were more likely to locate the water sources on private land. We installed 1.9 sources per village on public land when project staff made decisions, and 1.3 when communities took decisions. A significantly smaller number of individuals contributed money towards the water sources in the communities which took decisions without any imposed rules (5 individuals per village), when compared to the other two models (9 individuals).

Recent experimental studies have explored several aspects of 'participation'. One influential group of experimental studies examine the impact of a participatory or 'community-driven' development project compared to a control group which does not receive any intervention (Fearon, Humphreys, & Weinstein, 2009, 2011; Humphreys, de la Sierra, & van der Windt, 2012; Casey, Glennerster, & Miguel, 2012)[4]. The most closely related studies to this one explore variations in

[4]Another related group of experimental studies focuses on varying requirements for participation of women in

how participatory decision-making processes are implemented in projects to provide a local public good, conditional on implementing some kind of participatory decision-making approach: Olken (2010) compares decisions taken at representative-based meetings to those taken by direct election-based plebiscites and Beath, Christia, and Enikolopov (2013) compare decisions taken by secret ballot referenda to those taken at consultation meetings. Our study differs from these two studies in three ways. First, while these studies infer that participation in decision-making does influence decisions taken or other outcomes, since changes in the participatory process alter these outcomes, they do not directly measure the effect of introducing participation in decision-making itself. By including a treatment group in which the project is implemented, under otherwise identical conditions, without community participation in decision-making, we are able to measure the effect of introducing community participation in decision-making. Second, the two participatory decision-making processes we compare differ from those that these studies consider; neither decision-making by consensus nor decision-making without any imposed rules have previously been explored. Finally, the preceding studies have so far only reported results on how changing the participatory process alters the decisions taken, while we are also able to report data on the project impacts[5].

Our results confirm that involving communities in decision-making can lead to greater project impacts in terms of number of projects successfully completed and changes in reported access to safe drinking water. However, the results also suggest that devolving decision-making authority to the community without measures to avoid co-option of the decision-making process by influential groups or individuals can lead to an increased incidence of elite capture. In our case, the number of safe water sources constructed increases without any reported increase in access to safe drinking water.

The paper is structured as follows. Section 2 describes the setting, the experimental design and the data; section 3 describes the results, and section 4 concludes.

a particular decision-making process (Humphreys et al., 2012; Casey et al., 2012) or institution (Chattopadhyay & Duflo, 2004). Other related experiments examine changes in incentives to participate in school monitoring committes (Banerjee, Banerji, Duflo, Glennerster, & Khemani, 2010) or changes to the institutional structure of those committees (Pradhan et al., 2014); participation in monitoring of road construction projects (Olken, 2007) and public health care providers (Björkman & Svensson, 2009); participation in project targeting for household-level interventions (Alatas et al., 2013); and dispute resolution training to improve informal institutions (Blattman, Hartman, & Blair, 2014).

[5] Olken (2010) reports results on decisions taken, consistency with preferences of different groups within the community, knowledge about the project, and satisfaction with the project processes; Beath et al. (2013) report results on project satisfaction, and consistency of decisions with ex-ante preferences.

2 Setting, Experimental Design and Data

2.1 Arsenic Pollution Problem in Bangladesh

The context for this study is the arsenic contamination problem in rural Bangladesh. Arsenic occurs naturally in shallow groundwater in the delta region.

In the 1970s and early 1980s, many international agencies promoted the use of groundwater — water from wells — as a safer alternative to surface water — collected from ponds or rivers — which is often contaminated by pathogens. At the time, noone had realized that groundwater in the region sometimes has naturally occurring high concentrations of arsenic. Arsenic contamination is not readily detectable in water, and symptoms of arsenic poisoning only appear after years of exposure and accumulation in the body. Information about high concentrations of arsenic in tubewells emerged only in the mid-1990s. By that time, the damage was done; the resulting epidemic of diseases associated with arsenic exposure has been described as 'the largest mass poisoning of a population in history' (Smith, Lingas, & Rahman, 2000). In 2008, when this project began, UNICEF estimated that 20 million people were still using water from wells with arsenic concentrations above the Bangladeshi standard, which is itself five times higher than the WHO standard (UNICEF, 2008).

Creating access to safe drinking water in the presence of arsenic contamination presents a problem of providing a local public good. The great majority of drinking water sources in Bangladesh are privately owned, including almost all tubewells that have high concentrations of arsenic. Technologies to provide water with low concentrations of arsenic are considerably more expensive, and entail high fixed costs. Only the richest households can afford to purchase these sources themselves. For most households, they must be provided at the community level, at which high fixed costs can be shared among many people. As a result, communities who wish to improve access to safe drinking water must typically solve a collective action problem.

Several technologies are available to provide arsenic-safe drinking water, of which deep tubewells are the most common in rural Bangladesh. Deep tubewells draw water from deep aquifers (approximately 700-800 feet below ground level) that have low concentrations of arsenic. Standard deep tubewells are relatively expensive to install, but easy to use and maintain, and replacement parts are readily available. In some areas, arsenic safe water is available at lesser depths of approximately

300-400 feet. In these areas, shallow tubewells can provide arsenic-safe drinking water, at a lower installation than deep tubewells. Shallow tubewells are otherwise very similar to deep tubewells in terms of functionality, maintenance requirements and ease of repair.[6] In some areas, there is considerable seasonal variation in water pressure in the aquifer and standard deep tubewells may not provide year-round access to safe drinking water. An alternative design — the deep-set tubewell — can provide year-round access to safe drinking water in these areas. The pumping mechanism in the deep-set tubewell is installed below the surface of the ground, as opposed to on the surface in the standard design. This means that the deep-set tubewell is more expensive and more difficult to repair in case of failure than the standard deep tubewell, but it is equally convenient and easy to use.

In some areas, there is no accessible arsenic-safe aquifer – for example, where an intermediate layer of rock cannot be penetrated using local drilling techniques – and therefore it is not feasible to install tubewells. An alternative technology is the arsenic iron removal plant (AIRP). AIRPs remove arsenic by oxidation and filtration. They are more expensive, larger and significantly more difficult to operate and maintain than tubewells, and our experience suggested that communities strongly preferred tubewells.[7] As a result, we will throughout the paper report treatment effects by the type of feasible technology – AIRPs or tubewells – as well as the overall treatment effect.

2.2 Experimental Design

The project intervention consisted of a package of technical advice and subsidies for the provision of up to three safe drinking water sources per community. We carried out the interventions between 2008 and 2011, in partnership with a Bangladeshi non-governmental organization (NGO), NGO Forum for Public Health. NGO Forum for Public Health is a well-established actor in the water and sanitation sector with more than 30 years experience in the field.

We conducted our study in communities located in two *upazilas* (subdistricts): Gopalganj, about 60 miles southwest of Dhaka, and Matlab, about 30 miles southeast of Dhaka. We focused

[6] During the study implementation period, information emerged about a problem of manganese contamination in shallow tubewells. As a result, we replaced shallow tubewells we had already installed free of charge with alternative technologies, if they tested positive for manganese.

[7] Where tubewells were not feasible, we also offered communities the opportunity to install rainwater harvesting systems or a pond sand filter, but since no community selected either of these options, we do not describe them further in the paper. Both technologies have limitations with respect to tubewells or AIRPs

on these sites because of the severity of the arsenic contamination problem in the area more than 80% of pre-existing tubewells were arsenic contaminated and because the sites had not yet received other interventions to address the problem. We randomly selected 250 villages with greater than 75% arsenic contamination,[8] equally split between the two upazilas, and ranging in size from a minimum of 7 households to a maximum of 1103, with the median size 170 households.[9] Before interventions began, we carried out an information campaign about the arsenic problem, to ensure that all villages were initially equally well informed about the arsenic problem.

In the villages we studied, communities generally relied on tubewells to obtain drinking water at baseline. 97% of houses reported using a tubewell for their main source of drinking water[10]. The majority of the remaining households used surface water i.e. water drawn from ponds, rivers or canals, which is highly likely to be contaminated by bacteria. Of the households using tubewells, 48% reported that the tubewell was arsenic safe, 42% reported that it was unsafe; and 9% did not know whether it was safe or not. 25% of households used their own tubewell, but only 15% of these tubewells were arsenic-safe; 21% of households used another family member's tubewell, and 37% of these tubewells were safe; 17% of households used another private tubewell and 57% of these tubewells were safe; and 33% of households used a communal source, of which 72% were safe.

The intervention we studied was a project designed to improve access to safe drinking water by providing arsenic safe sources of drinking water. Research assistants first organized a community meeting to discuss the project, involving the community to a greater or less extent in planning the meeting, depending on the type of decision-making structure studied. At the community meeting, the research assistants discussed the need for access to safe drinking water and explained the terms of the intervention. In particular, they informed the village that funds were available to fund up to 3 safe sources of drinking water, conditional on the community contributing towards the cost of construction, and how the relevant decisions were to be taken. After the meeting, the community had up to twelve weeks to raise the community contribution. If the contribution was met for a given location, the research assistants arranged for contractors to come to the village and construct the water source in the agreed location. The research assistants or the communities

[8]We excluded urban and peri-urban areas in Matlab. In Matlab, five villages were included in the study despite having arsenic contaminations of between 65 and 75%.

[9]Data on arsenic contamination of pre-existing tubewells and village size was drawn from the Bangladesh Arsenic Mitigation Water Supply Project.

[10]Estimates in this section population-weighted and calculated for the control and treated villages.

chose a maintenance committee, who received a basic toolkit and maintenance manual, and who would then be responsible for maintenance and repair of the installed source.

Of the 250 villages studied, we assigned 100 to a control group who did not receive the intervention. 126 villages received the intervention. We initially assigned a further 24 villages to receive the intervention who eventually did not receive the intervention, due to changes in the costs of providing safe water sources over the course of the project. We originally assigned one other village to treatment, but project staff determined before the project began that there were no feasible available technologies to provide safe drinking water in the community, because no arsenic safe aquifer was accessible, and arsenic concentrations in the shallow groundwater were too high for removal with an AIRP. There was one other village in which we determined after we began the intervention that there were no feasible available technologies to provide safe drinking water.

The original protocol for selection of treated villages was random, which should have resulted in treatment and control groups which were comparable at baseline. However, we later established that the project director at the time, who was later removed from the project for unrelated reasons, did not follow the original protocol when he implemented the division of the villages into control and study villages, and he included all villages in the southern area of Matlab in the treatment group. Villages in South Matlab have much lower access to safe drinking water than the average village in the sample, meaning that overall the treated group had significantly lower access to safe drinking water at baseline than the control group.

Table 1 confirms that this resulted in statistically significant differences between control and treatment groups. Treated villages had reported lower access to safe drinking water, and were less likely to have changed their source of drinking water because of the arsenic contamination problem in the last five years. In Table 1, we show baseline summary statistics and randomization checks for villages by treatment status. The table shows the mean and standard errors for a selection of baseline variables which measure baseline access to safe drinking water, factors that might influence the ease of providing safe drinking water, and community-level variables that might influence the likelihood of a successful collective action. In column 2), we test whether the difference in means between treated and control villages is statistically significant. The p-values are derived from Ordinary Least Squares (OLS) regressions with the following structure:

$$Y_{i,v} = \alpha + \beta I_{treated,v} + \epsilon_{i,v} \tag{1}$$

where $Y_{i,v}$ is the value of a variable in household i in village v and $I_{treated,v}$ is an indicator which is one if village v was treated and zero if village v was not treated. If the treatment was randomly assigned, the coefficient β should be zero as assignment to treatment should not be correlated with any baseline characteristics of the village. The p-values test whether the coefficient is equal to zero.

Since treatment was assigned at the village level, but we collected data at the household level, it is important to account for within-village correlation in variables. Within-village correlation implies that it is more likely that differences between mean outcomes in treated and control villages arise due to chance, than if we had been able to assign treatment at the household level. In order to ensure that the statistical analyses we carry out make the correct inference about whether or not a result is likely to be due to chance or not, we follow Angrist and Pischke (2009) and cluster standard errors at the village level.

Columns 3) to 5) of Table 1 show that we can correct for the bias induced by the failure of randomization by three methods. First, we can drop South Matlab from the sample. Second, we can create a synthetic treatment variable generated at random in South Matlab, and equal to the treatment variable elsewhere [11]. This synthetic variable re-assigns a fraction of the villages in each treatment group in South Matlab to control. Third, we can use this synthetic treatment variable to instrument for treatment. In columns 3) to 5), we report the difference in means between treated and control villages under these three approaches, after accounting for the different proportions of treated villages in Gopalganj and Matlab, because differences in treatment and control groups otherwise reflect differences between these areas. We estimate the difference in means using an equation similar to Equation 1, including indicators for Gopalganj and South Matlab. In each case we show that no significant differences remain between treatment and control populations.

Since the non-random selection of treatment villages in South Matlab may have introduced bias into our estimates of treatment effects, we therefore report both OLS results and results where treatment is instrumented using the synthetic treatment variable (the IV results).

[11] Ideally, we would have used the original random assignment to treatment rather than this synthetic alternative but we have not been able to recover the initial, randomly assigned treatment lists.

Decision-making structures

Project staff implemented the intervention under one of three decision-making structures. The necessary decisions included if, how and where to install; and how to manage, each safe drinking water source. In all cases, project staff ensured that all decisions made were technically appropriate. Table 2 summarizes the main features of the different decision-making structures. We describe the three models in more detail in the following paragraphs.

The decision-making structures included one non-participatory structure, the Top-Down model (TD). Under this model, project staff took all project decisions, after an extended (typically 2-day) period of information gathering. The information gathering process consisted of participatory mapping of the village with members of the community, focusing on the locations of households and safe and unsafe sources of drinking water, cross-checking information with various community members. Project staff then proposed sites for safe drinking water sources, prioritizing locations with the highest density of households not already served by safe drinking water sources, choosing public locations wherever possible, and convenient locations where no suitable public land was available. Staff then organized and publicized a community meeting at which they presented the proposed locations. This model was designed to approximate the 'traditional' approach to decision-making about local public goods in which decisions are taken by a centralized organization, such as local government or an NGO.

The decision-making structures also included two participatory structures, in which decision-making authority was devolved to the community. Under the 'pure' Community Participation (CP) model, project staff visited the community to arrange a meeting at a site and time of the community's choosing. At the meeting, project staff explained the project rules and announced that they would return to the village after a few days to find out whether they wanted to participate in the project, and if so, which sites they had chosen. Sites that were not technically appropriate were rejected, but otherwise the community's decisions were final, conditional on raising the community contribution. We did not directly observe the decision-making process used, but communities reported to us that they took these decisions in a variety of ways including open meetings (sometimes but not always including women), meetings at a mosque, or closed-door meetings of village elites. This model was designed to approximate the way in which some organizations implement community participation in practice, avoiding interference with a community's internal hierarchies and decision-making processes .

Under the second participatory decision-making structure, the NGO-Facilitated Community Participation model (NGO), we imposed rules about how decisions should be taken. Project staff initially organized a series of separate small group meetings with men and women who the community identified as poor and non-poor. At these small group meetings, project staff explained the project rules and emphasised the right all individuals would have to participate in the decision-making process and benefit from the interventions. These meetings were followed by a community meeting, at which both men and women, and poor and non-poor, had to be represented. The community proposed and selected water source locations by unanimous consensus at the meeting, in the presence of project staff and with their active facilitation. If the community could not reach a consensus at the first meeting, a second and in some cases subsequent meetings were organized. This model was designed to approximate the way in which other organizations implement community participation, with project staff playing a strong facilitatory role, and rules imposed that are intended to reduce the likelihood that influential groups or individuals co-opt the decision-making process.

Before installing a safe drinking water source, we required the community to contribute between 10% and 20% of its cost, depending on the technology installed. Table 3 shows the cost of installing each of these technologies and the community contribution that we required. The difference in required community contributions reflects the difference in cost of the selected technology. We also scaled the community contribution so that the subsidy could be either concentrated on one water source or spread between up to three water sources. The price per water source therefore increased as more water sources were installed in the village. Budget constraints meant that when the best feasible technology was one of the more expensive alternatives, we were only able to offer up to two water sources.

After the initial decision-making process, project staff gave the communities up to twelve weeks to raise the funds for the community contribution. Construction of the safe drinking water sources began as soon as the community had raised their contribution. If after twelve weeks the community had not raised their contribution, construction of the safe drinking water sources did not go ahead. We initially intended the decision-making structures to apply to decisions about who contributed to the community contribution, but this proved impossible to enforce. However, project staff did propose a list of contributors at the Top Down model meetings, and communities did agree a list of contributors at the NGO-Facilitated Community Participation meeting.

We randomly assigned the decision-making structures to the communities who received the intervention. Of the 126 treated villages, we initially assigned 42 to each decision-making model. We replaced the village in which we determined before beginning the project that there was no feasible safe drinking water technology with another village, randomly drawn from the villages which we had initially assigned to treatment but in which we had not carried out the intervention due to budget constraints. As a result 43 villages were assigned to the Top-Down model.

Table 4 shows that the villages assigned to each decision-making model were comparable at baseline to the villages assigned to the other decision-making models. We test whether the difference in variable means between villages in which the project was implemented under a given decision-making structure and the remainder of the treated villages is statistically different from zero. The p-values in the table are therefore derived from OLS regressions similar in structure to Equation 1 but the indicator $I_{m,v}$ is one if village v received treatment under decision-making structure m, and zero otherwise:

$$Y_{i,v} = \alpha + \Sigma\beta_m I_{m,v} + \epsilon_{i,v} \tag{2}$$

Only the treated villages are included in the regressions in Table 4. We do not use the control group for comparison in this case because the results in Table 1 already confirm that the treated villages are not directly comparable to the control villages.

We compare 15 variables across the 3 decision-making structures, resulting in a total of 45 tests. In 43 of these tests we fail to reject at the 10% level the null hypothesis that there is no difference in means between groups treated under one decision-making structure and the other treated villages. In 2 tests we find statistically significant differences between the mean of a variable in villages treated under one decision-making structure and in the remaining treated villages. One test rejects this hypothesis at the 5% level, and one at the 10% level. This is consistent with what we would expect due to chance. From these checks we conclude that there is no evidence to suspect that assignment to model, conditional on treatment, was not random, as required by the project protocol.

The same project staff – one team in Gopalganj and one team in Matlab – implemented the project under all three decision-making structures. We implemented the intervention in cycles during which project staff would complete the entire process from meeting organization to water source installation for a group of villages, where the villages were grouped geographically for ease of logistics. The project was initially implemented in 114 villages in 6 cycles across both upazilas. We later added an additional 12 villages in Gopalganj when funds became available, in a 7th cycle.

Government policy had changed by the time we carried out the 7th cycle, and community expectations that the government would provide free tubewells may have increased. We installed fewer safe water sources under the 7th cycle, but the number installed is not significantly less than under the first 6 cycles in Gopalganj, once we account for the feasible technology.

2.3 Data Description

We carried out a baseline survey in 2007 in 40 households in each of the 250 villages, sampled randomly from census lists . We surveyed a total of 9797 households, as in some very small villages there were fewer than 40 households. The baseline questionnaire included standard components of a household survey with a special focus on social networks and social capital, and full details on water use behavior. We also collected village-level information from focus groups.

We encountered significant problems with the data entry process after the baseline survey. First, some of the individuals employed to enter the data in spreadsheets copied and pasted entire villages of data, changing names and other identifiers to conceal what they had done. Data checking revealed this problem by chance several months after data collection and entry had been carried out. When we discovered this problem, we checked extensively for additional incidences and had the missing data re-entered. Second, by the time we discovered this problem, termites had unfortunately attacked the stored questionnaires, and destroyed a small percentage of the questionnaires. As result, we are missing baseline data from 140 households from control and treated villages, since enumerators did not initially enter the data correctly and termites then destroyed the hard copy of the questionnaires. We do not however have any reason to think that there was any systematic pattern to either the false data entry or the losses to termites, so the remaining baseline data should still represent a randomly selected sample of the baseline population.

We carried out follow-up surveys in control and treated villages in 2010 and 2011 after we carried out the safe water intervention, interviewing the same households that we interviewed for the baseline survey. We did not carry out follow-up surveys in the 24 villages which were initially assigned to treatment but in which we did not carry out the intervention. We therefore attempted to resurvey 8,890 households from the original panel, of which we successfully re-surveyed 8630 households, representing an average attrition rate of 2.9%. The attrition rates broken down by treatment group are as follows: 2.7% in control villages; 3.1% in treated villages. Among the treated villages, attrition rates were 2.6% in NGO-Facilitated Community Participation villages; 3.2% in Community Participation villages; and 3.4% in Top-Down villages. The attrition rates in treated groups and sub-groups were not statistically different from the control group, or from each other.

We also carried out follow-up surveys in 1424 additional households in treated villages, to bring the minimum survey coverage up to 15% of households in all treated villages (based on census data). The additional households were randomly selected from the remaining households on the census lists who had not been surveyed at baseline. Extending the survey coverage in this way was intended to ensure that the survey captured the effects of the intervention in larger villages, where the three safe drinking water sources constructed were unlikely to serve the entire community. However, the data from these additional households is inconsistent with the data collected from the panel households, and we have established that there were violations of the sampling protocol; in particular, in some villages, some of these additional households were sampled from the neighbourhood of the installed water sources, rather than from the census lists. As a result, we do not use this data in this analysis.

We also collated data on the numbers and types of safe drinking water sources installed, and project staff kept detailed records of the implementation process, including the number of contributors in each community and the time taken to raise the community contribution. We also carried out focus group discussions in treatment villages to obtain qualitative information about why the project was successful in some communities and not in others.

3 Results

We first show how attendance at decision-making meetings varied by decision-making model to demonstrate that the NGO-Facilitated Community Participation was marginally more successful in including more people in the decision-making process, and in including a more diverse range of people. We then report the project outcomes in the study villages. Projects implemented under the participatory decision-making models were more successful in terms of installing safe water sources, but the differences between models are not statistically significant. We were far more successful in installing safe water sources in villages where tubewells were feasible, than in villages where only AIRPs were feasible.

Projects implemented under the Top Down decision-making model were however much more successful in installing safe water sources in public places, although whether an installation site is public or not is not necessarily a good predictor of how well-used the water source will be. Finally, we show that in villages implemented under the Community Participation model, fewer households contributed towards the cost of installation.

We then report the average treatment effect in terms of changes in reported access to safe drinking water for all villages. We show that overall, the project led to a 16% increase in reported access to safe drinking water. The treatment effect was higher in villages where tubewells were feasible (18%), and zero in villages where only AIRPs were feasible. We then show that the treatment effect was substantially higher in villages where the project was implemented under the NGO-Facilitated Community Participation model, and that the differences are significant when we exclude the villages in which only AIRPs were feasible; given that the average treatment effect is zero in these villages, including these villages is not informative with respect to a comparison between decision-making models. Finally, we present some robustness checks on these main results.

3.1 Participation

Table 5 shows the recorded numbers and characteristics of individuals attending the main community meetings. Overall, the mean number of participants was 30.7, with 27% of meeting attendees female, 42% of low socioeconomic status (as recorded by project staff) and 60% with less than secondary education.

More people, from a more diverse range of groups, attended meetings held under the NGO-Facilitated Community Participation decision-making model than meetings held under the other two decision-making structures. The number of participants was somewhat higher in NGO-Facilitated Community Participation model meetings (33.6), and lowest in the Top Down model meetings (28.5), but the differences between models are not statistically significant. The NGO-Facilitated Community Participation model meetings were the most diverse in representing different groups, with the highest percentage of female participants and the highest percentage of participants with less than secondary education. The Top Down meetings were the least diverse, with the lowest percentage of female participants, and the lowest participation of participants with low socio-economic status. However, again, not all the differences between models are statistically significant. Nonetheless, to the extent that attendance at meetings really reflects *de facto* participation in decision-making, the NGO-Facilitated Community Participation model was more successful in involving a larger number and wider range of community members in decision-making.

3.2 Project Outcomes

Table 6 shows how the decision-making model influenced project outcomes in the treated villages. On average, we installed 2.14 safe water sources in the treated villages. If we had installed all technically feasible water sources given our project rules, we would have installed an average of 2.75 safe water sources, based on installing three sources per village in most cases, and two sources per village where only a more expensive technology was feasible.

We offered communities the choice between all technically appropriate technologies to provide safe drinking water, given local hydrogeological conditions. In Gopalganj, we carried out the intervention in 70 villages. In 16 villages, AIRPs were the only feasible technology. In two villages, no treatment was feasible, as there was a layer of impenetrable rock, and shallow groundwater was too strongly contaminated with arsenic and iron for removal with an AIRP. In Matlab, tubewells were feasible in all villages.

A clear preference gradient between the available technologies emerged. People chose shallow tubewells wherever possible, followed by standard deep tubewells and deepset tubewells. The three types of tubewell are comparable in ease of use and maintenance, but increase in cost with depth and design complexity. AIRPs were the least preferred option by a wide margin. There were 16 villages, all in Gopalganj, where the only type of water source that could be installed was an AIRP, meaning that we could have installed a total of 32 AIRPs. We were only successful in installing 5 AIRPs during the course of the project, a success rate of approximately 16%. In comparison, in the remaining villages in Gopalganj — in which tubewells were feasible — we installed 79% of the maximum number of wells we could have installed under our project rules. The reasons given by the communities for rejection of the AIRPs were that they took up too much space, required too much work to operate and maintain, and were not perceived to be reliable or trustworthy. When we consider only the villages in which tubewells were feasible, the average number of water sources constructed rises to 2.45 out of a maximum possible 2.85.

The rejection of AIRPs did not seem to be a direct function of the price of the technology. However, in Matlab, in the 10 villages where only deep-set tubewells could be installed (which are comparable in price to AIRPs, and for which we required the same level of community contribution), we installed on average 90% of the maximum feasible number of water sources, compared to an average of 89% in all other villages in Matlab (where either deep tubewells or shallow tubewells were feasible). However, we cannot comment on what would have happened if we had offered AIRPs at a lower price.

We installed 10% more water sources in the villages in which communities participated in decision-making than in the villages in which project staff took decisions, as shown in column 1). Installing more water sources is one measure of success of the project, but it may not translate into increased access to safe drinking water if the sources are not fully accessible to the community. However, the differences are not statistically significant. In Table 6, we assess whether differences in project outcomes across models are statistically significant using OLS regression for the following equation:

$$Y_v = \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_v \tag{3}$$

We then test pairwise equality of the coefficients β_{NGO}, β_{CP} and β_{TD}. The differences between the number of water sources installed under the different decision-making models are attenuated further in both magnitude and significance when we consider only the villages in which tubewells were feasible.

We installed more water sources in public spaces, as recorded by our project staff, under the non-participatory Top-Down model. Public spaces were defined to include communal land, open spaces, areas beside roads, and institutions such as mosques or schools, as opposed to privately owned land. Under the Top-Down model, project staff had a specific mandate to install water sources in public places. The differences are strongly significant with respect to both the participatory decision-making models. Water sources installed in public places may be accessible to a larger number of people. However, space that is appropriate for water source construction is quite strongly constrained in villages in this region, and the most convenient location for a water source may not necessarily be located on public land. This is primarily because land that is not vulnerable to flooding is relatively scarce, and safe water sources cannot be installed on land that is vulnerable to flooding because of potential contamination.

Fewer people contributed to raising the community contribution in the unregulated Community Participatory model than under the other two models. The difference is significant with respect to both the other two decision-making models. A small number of contributors may be efficient, as some community members will have a greater ability to contribute than others. However, it may also be indicative of a high degree of influence over the decision-making procedure, which may not be efficient if used to co-opt project benefits for private use.

Overall, the number of contributors was relatively low in all cases, considering that the median village size was 170 households. In villages where we successfully installed at least one safe water source, the mean number of contributors per water source installed was 5.1 in NGO villages, 2.3 in CP villages and 4.0 in TD villages. There was only one contributor per safe water source installed in 34% of the NGO villages, 56% of the CP villages and 45 % of the TD villages.

3.3 Reported Project Impact

We primarily measure access to safe drinking water based on an outcome variable which measures whether or not the household reports using safe drinking water. The indicator is based on the source of water that the household identifies as being its most important source of water for drinking and cooking. The indicator for reporting use of safe drinking water is constructed as being equal to one where the household reports using a source of drinking water that is safe from both bacterial and arsenic contamination, and zero when they report that the source is unsafe, if they don't know whether it is safe or not, or if it is a source that is vulnerable to bacterial contamination e.g. a dug well or surface water. Further details regarding the construction of this variable is included in Appendix A.

We report the average overall treatment, relative to the full control group, but we also break down the treatment effect by whether tubewells or only AIRPs were feasible. There is strong spatial correlation between locations where only AIRPs are feasible, reflecting the extent of the rock layer overlaying the deep aquifer. Since other village level characteristics are also spatially correlated, there are as a result some differences on baseline characteristics between villages in which tubewells were feasible and villages in which only AIRPs were feasible in Gopalganj.

When we report effects for villages in which a specific technology was feasible, we use a matched control group, because we do not observe which technologies are feasible in the control villages. Using a matched control group removes these differences (see Appendix Table B1). However, in robustness checks, we show that the results for tubewell villages are not sensitive to using a different matched control group, or to simply using the full set of control villages. The construction of the matched control group exploits spatial correlation in the location of villages in which AIRPS were the only feasible technology. Details of the construction of the matched control group are given in Appendix A.

Average treatment effect

In Table 7 we show reported access to safe drinking water at baseline (Panel A) and follow-up (Panel B), and the resultant change in access (Panel C). We show results for all villages in columns 1) to 3); in all villages in which tubewells were feasible in columns 4) to 6); and in those villages where only AIRPs were feasible in column 7).

Columns 1) and 4) show the OLS results, which as previously discussed show that treated villages have worse access to safe drinking water at baseline. Columns 2) and 5) show the results dropping South Matlab, where the randomized assignment to treatment was not correctly implemented. Columns 3) and 6) show IV results, using a synthetic assignment to treatment variable in South Matlab. EIther of these approaches removes the baseline differences in reported access to safe drinking water. In Gopalganj, there were no problems with random assignment to treatment, so we only report one set of results.

To estimate the results in Panels A) and B), we use data from all panel households and estimate the following equation:

$$_{i,t} = \alpha + \beta I_{treated,v} + \epsilon_{i,t} \tag{4}$$

where i is a household in village v and $Y_{i,t}$ is the access to safe drinking water at baseline or followup.

Panel A) shows baseline access to safe drinking water. This panel repeat the comparisons shown in Table 1, and demonstrates the validity of the approaches used for compensating for the failure of random assignment. Panel B) shows the follow-up comparisons. As a result of the initial differences between treatment and control villages, the differences between follow-up and control villages are not statistically significant at follow-up in the OLS analysis, but are clearly evident when we correct for the failure of random assignment.

To estimate the change in access to safe drinking water between baseline and followup (Panel C), we estimate a first difference equation for all households for which we have both baseline and follow-up data, as follows:

$$\Delta Y_i = Y_{if} - Y_{ib} = \alpha + \beta I_{treated,v} + \epsilon_i \tag{5}$$

where ΔY_i is the change in access to safe drinking water between baseline and followup. With two time periods, the first difference analysis is directly equivalent to including household fixed effects. As before, we cluster standard errors at the village level to account for within-village correlation in outcomes.

Panel C) shows the change in reported access to safe drinking water. The OLS and IV results are almost identical, suggesting that although assignment to treatment was not random in all areas, it was not correlated with trends in reported access to safe drinking water. The estimated average treatment effect is 16% overall, and 18% in villages in which tubewells were feasible. There is no treatment effect in the AIRP villages.

For these results and the remainder of the results in this section, we use survey weights which ensure that each village counts equally in the analysis. Where part of the data for a village was lost through the baseline data entry problems, the baseline weights compensate for these losses, as there is no reason to think that the lost data introduces any bias to the estimates of a variable in the village. We do not introduce compensatory weights for migration, but attrition rates were low overall, so this is unlikely to influence the results.

Treatment effect by decision-making model

Table 8 repeats the analyses shown in Table 7, with the estimated effects broken down by the decision-making model under which we implemented the project. Once again, Panel A) shows baseline access to safe drinking water, Panel B) shows access at follow-up, and Panel C) shows the change in access. Columns 1) to 3) show the results in all villages; columns 4) to 6) show the results in villages where tubewells were feasible; and column 7) shows the results where only AIRPs were feasible. In columns 2) and 5), we correct for failure of random assignment to treatment by dropping South Matlab; and in columns 3) and 6) we correct by instrumenting for the model assigned by the interaction between an indicator for the implemented model, and the synthetic treatment variable.

In Panel A) and B), we use equations with the following structure:

$$Y_{i,t} = \alpha + \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_{i,t} \tag{6}$$

where $Y_{i,t}$ is reported access to safe drinking water at baseline or follow-up in household i in village v, and I is an indicator whether the village was treated under a given decision-making structure.

Panel A) shows that villages treated under each decision-making model differ from the control villages (as a result of the failure of random assignment to treatment), but that the villages treated under each model are comparable to each other. Columns 2), 3), 5) and 6) also confirm that the strategies for correcting for the failure for random assignment to treatment are also effective for removing significant baseline differences between villages treated under a given model and the control villages as a whole. This is consistent with the results shown in Table 4, in which we showed that conditional on treatment, villages assigned to different decision-making models were comparable on baseline statistics. However, note that the magnitude of differences between treatment groups is quite large with respect to the treatment effects we estimate. In particular, 36% of households in NGO Facilitated Community Participation villages report having access to safe drinking water at baseline in comparison to 41% in Community Participation villages and 44% in Top Down villages.

Panel B) shows that, in the specifications when we correct for the failure of random assignment to treatment, there are significant treatment effects under all decision-making models. The estimated effects using only follow-up data are largest for the NGO Facilitated Community Participation, but the differences across models are generally small. However, note that the pattern of access to safe drinking water is reversed: villages in which the project was implemented under NGO Facilitated Community Participation model have the lowest access to safe drinking water at baseline, and the highest access to safe drinking water at follow-up.

In Panel C), we show the results for the change in reported access to safe drinking water. We find that the reported increase in safe drinking water is greatest in NGO model villages, and the reported increase in safe drinking water is almost exactly equivalent in TD and CP model villages. We estimate a first difference regression of the change in reported access to safe drinking water using the following equation:

$$\Delta Y_i = Y_{i,f} - Y_{i,b} = \alpha + \beta_{NGO} I_{NGO,v} + \beta_{CP} I_{CP,v} + \beta_{TD} I_{TD,v} + \epsilon_i \tag{7}$$

The estimated coefficients are extremely consistent, with the estimated increase in access to safe drinking water between 21% and 22% in NGO-Facilitated Community Participation model villages (24% and 25% in tubewell villages); between 13% and 14% in Community Participation model villages (between 13% and 15% in tubewell villages) and between 11% and 13% in Top Down model villages (between 12% and 15% in tubewell villages). None of the models shows any significant treatment effect in AIRP villages.

The size of the effect is economically quite important, as the treatment effect almost doubles. However, the differences between models are just below the threshold of statistical significance when all villages are considered. The difference in size of the treatment effect between the NGO-Facilitated Community Participation model villages and the other treated villages is statistically significant when we exclude the villages in which only AIRPs were feasible. Since the treatment effect was zero in these villages, including these villages reduces all the estimated model-specific treatment effects, making it more difficult to distinguish between them, and introduces noise that is not informative with respect to a comparison between the decision-making models.

We do not include data from the additional households that we surveyed at follow-up, because of inconsistencies between the additional households and the panel households, and because of concerns that the sampling protocol may not have been implemented consistently. Including data from the additional households surveyed at followup increases the magnitude and the statistical significance of the difference between the NGO model villages and the other treated villages.

3.3.1 Robustness Checks

In Table 9, we show the effects of changing the main specification on the estimates of the size of the treatment effects under the different decision-making models. Unless specified, we focus only on the tubewell villages, where the differences between models are statistically significant. For brevity, we only report the p-values of the tests of interest, given the main results: whether the results from the NGO-Facilitated Community Participation model villages are equivalent to the results from the other villages.

In Column 1), we show that the results are similar when we consider only the treated villages, and do not include the control villages in the analysis. The results are given as OLS as within the treated villages, assignment to model was random.

In columns 2) and 3) we show the results by upazila. The treatment effects are larger in Gopalganj, and the differences between models are more pronounced. However, the pattern of results is also consistent for Matlab in that the increase in access to safe drinking water is largest in the villages treated under the NGO-Facilitated Community Participation model.

In columns 4) and 5), we show that the results are not sensitive to how the matched control group is constructed. Column 4) shows that an alternative matched control group, generated by assigning villages in Gopalganj to AIRP and tubewell-matched groups at random, given the probability of AIRP/tubewell feasibility in their neighbourhood, yields almost identical results.

Column 5) shows that using the full control group also yields similar estimates.

In columns 6) to 8), we include additional controls. We include the interactions between the control variables and the treatment variables, since otherwise including control variables results in bias in the estimated treatment effect (Freedman, 2008)[12]. In column 6), we use the full set of treated villages and control for the best available technology. We do not include the control villages as we do not directly observe the best available technology in these villages. The coefficients are very similar.

In column 7), we include a quadratic function of village size, and its interactions with the treatment indicators. Allowing for heterogeneity in the treatment effect by village size increases the size of the estimated effect in NGO-Facilitated Community Participation. This reflects substantial heterogeneity in the overall effect, and the effect by model, across village size. Figure 1 shows the heterogeneity in the overall treatment effect. The effect size decreases with village size, reflecting the fact that the intervention was limited to install a maximum of three safe water sources, regardless of village size, and possibly, the increasing difficulty of solving a collective action problem with a group of increasing size, as theory predicts (Olson, 1971). No treatment effect is detectable in villages with more than 500 households, although the number of villages in this group is small (8 treated villages and 10 control villages). Figure 2 shows the heterogeneity in the treatment size by decision-making model:, the effect sizes decreases for each decision-making model, with the effect size greatest under the NGO Facilitated Community-Participation model over the entire range of sizes at which treatment effects are observed. We use the full set of treated and control villages in this column, and column 8.

In column 8), we include a quadratic function in total household assets and its interactions with the treatment indicator. The results are similar to the main specification.

In columns 9) and 10) we report results using alternative measures of access to safe drinking water, described in detail in Appendix 4. In column 9, we report results using a measure of whether the source the household was using could be verified to be safe. To verify safety of a tubewell, enumerators inspected the tubewell that the household reported using if it was less than

[12]We include interactions with the raw treatment indicator, rather than the synthetic treatment indicator used for the first stage, as we are not clear which is the correct procedure in the context of an instrumental variables regression. However, the change in the results is minor if we instead control for interactions with the synthetic treatment variables.

5 minutes walk away, and recorded whether it was marked red (unsafe), green (safe) or unmarked. The increased treatment effects are larger across all models, partly because our intervention also increased the fraction of tubewells that were verifiably safe, and also because we did not collect this data from all villages at baseline. In column 10), we report results using a measure than uses the verified data when it is available, and the reported data when enumerators were not able to verify the safety of a source. The results are broadly consistent with the results which use only the reported measures.

4 Conclusions

This study has provided the first experimental evidence to support the claim that delegating decision-making authorities to communities in projects to provide local public goods can improve outcomes and increase reported impact. In villages where we implemented a project to provide safe sources of drinking water under a participatory decision-making structure (the NGO Facilitated Community Participation model), we installed a slightly larger number of safe drinking water sources (0.2 more sources) but obtained a 9% higher increase in access to safe drinking water, than under a non-participatory decision-making structure (the Top-Down model). Under this model, community members took key decisions by unanimous consensus at a community meeting with minimum representation requirements. These results are broadly consistent with evidence accumulated in the past through practitioner's experience and cross-sectional analysis, but this is the first time that experimental evidence has been available to test the hypothesis that participation in decision-making has a positive impact on the result of social programs.

However, the study also suggests that these benefits may not be realised if protective measures are not put in place to prevent the decision-making process from being co-opted by influential groups or individuals. Under the 'pure' Community Participation model, under which communities took decisions without imposed rules, we installed the same number of sources as under the NGO-Facilitated Community Participation model, but we obtained a 9% smaller increase in access to safe drinking water.

Since we did not test alternative strategies for preventing the decision-making process from co-option, we cannot comment as to whether the method used here (imposing the requirement that decisions be taken by unanimous consensus at a community meeting where all groups were represented and conducting small group meetings beforehand to raise awareness about the project objectives and the rights of all individuals to participate) was the most effective possible in the context. We also did not delegate technical decision-making authority to the community (our project staff determined the feasibility of any given technology and location) and therefore cannot determine whether the results would be the same or different if decision-making authority is delegated to the community over other types of decisions.

A potential weakness of our results is that we rely on reported data, and it is possible that participation in project decision-making may influence the way in which intended beneficiaries report project outcomes. We have also collected data on actual use of the installed water sources by monitoring their use directly using enumerator observations. This data is currently being analysed.

The role of the community contribution appears key in determining outcomes. The number of contributors is low over all. Those that can contribute towards the cost of the water source may have significant influence over the decision-making process. The number of contributors is lowest in the pure Community Participation villages, where we find suggestive evidence of a higher degree of elite capture. Anecdotally, project staff reported to us that in some Top-Down model villages where community groups failed to raise the community contribution, individuals volunteered to pay the community contribution, but only if the water source was installed on their private land.

The result of delegating decision-making authority to the community may vary substantially depending on the local context, for example depending on existing inequalities within the community or on the size and homogeneity of the group to which authority is delegated. We cannot determine whether the results of this study would be applicable in other contexts. The study would benefit from replication in different social and cultural contexts.

Bearing these caveats in mind, the results provide important experimental evidence regarding an influential policy recommendation, and suggest that careful consideration should be given to the structure of a participatory decision-making process, if the potential benefits are to be realized.

References

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., Purnamasari, R., & Wai-Poi, M. (2013, January). *Does elite capture matter? Local elites and targeted welfare programs in Indonesia*.

Angrist, J. D. & Pischke, J. (2009). *Mostly harmless econometrics*. Princeton University Press.

Banerjee, A., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*.

Beath, A., Christia, F., & Enikolopov, R. (2013, January). *Direct democracy and resource allocation: Experimental evidence from Afghanistan*. Centre for Economic and Financial Research at New Economic School Working Paper No 192.

Björkman, M. & Svensson, J. (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics*, *124*(2).

Blattman, C., Hartman, A. C., & Blair, R. A. (2014). How to promote order and property rights under weak rule of law? an experiment in changing dispute resolution behavior through community education. *American Political Science Review*, *108*(01).

Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *Quarterly Journal of Economics*, *127*(4).

Chattopadhyay, R. & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, *72*(5).

Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2009). Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *American Economic Review: Papers and Proceedings*, *99*(2).

Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2011, December). *Democratic institutions and collective action capacity: Results from a field experiment in post-conflict Liberia*.

Freedman, D. A. (2008). On regression adjustments to experimental data. *Applied Mathematics*, *40*, 180–193.

Fritzen, S. (2007). Can the design of community-driven development reduce the risk of elite capture? Evidence from Indonesia. *World Development*, *35*(8), 1359–1375.

Fung, A. & Wright, E. O. (2003). *Deepening democracy: Institutional innovations in empowered participatory governance*. Verso.

Hossain, M. A., Sengupta, M. K., Ahamed, S., Rahman, M. M., Mondal, D., Lodh, D., ... Chakraborti, D. (2005). Ineffectiveness and poor reliability of arsenic removal plants in West Bengal, India. *Environmnetal Science Technology*, *39*(11), 4300–4306.

Humphreys, M., de la Sierra, R. S., & van der Windt, P. (2012, June). *Social and economic impacts of Tuungane: Final report on the effects of a community driven reconstruction program in Eastern Democratic Republic of Congo*.

Isham, J., Narayan, D., & Pritchett, L. (1995). Does participation improve performance? Establishing causality with subjective data. *World Bank Economic Review*, *9*(2), 175–2000.

Khwaja, A. (2004). Is increasing community participation always a good thing? *Journal of the European Economic Association: Papers and Proceedings of the Eighteenth Annual Congress of the European Economic Association*, *2*(2/3), 427–436.

Khwaja, A. I. (2009). Can good projects succeed in bad communities? *Journal of Public Economics*, *93*.

Kleemeier, E. (2000). The impact of participation on sustainability: An analysis of the Malawi rural piped scheme program. *World Development*, *28*(5), 929–944.

Mansuri, G. & Rao, V. (2013). *Localizing development: Does participation work?* Washington, DC: World Bank.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *115*(2).

Olken, B. A. (2010). Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American Political Science Review*, *104*(2).

Olson, M. (1971). *The logic of collective action*. Harvard University Press.

Pradhan, M., Suryadarma, D., Beatty, A., Wong, M., Gaduh, A., Alisjahbana, A., & Artha, R. P. (2014). Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia. *American Economic Journal: Applied Economics*, *6*(2).

Rao, V. & Ibáñez, A. M. (2005). The social impact of social funds in Jamaica: A 'participatory econometric' analysis of targeting, collective action, and participation in community-driven development. *Journal of Development Studies*, *41*(5), 788–838.

Sara, J. & Katz, T. (1997). *Making rural water sustainable: Report on the impact of project rules*. UNDP – World Bank Water and Sanitation Program.

Smith, A., Lingas, E., & Rahman, M. (2000). Contamination of drinking water by arsenic in Bangladesh: A public health emergency. *Bulletin World Health Organization*, *78*, 1093–1103.

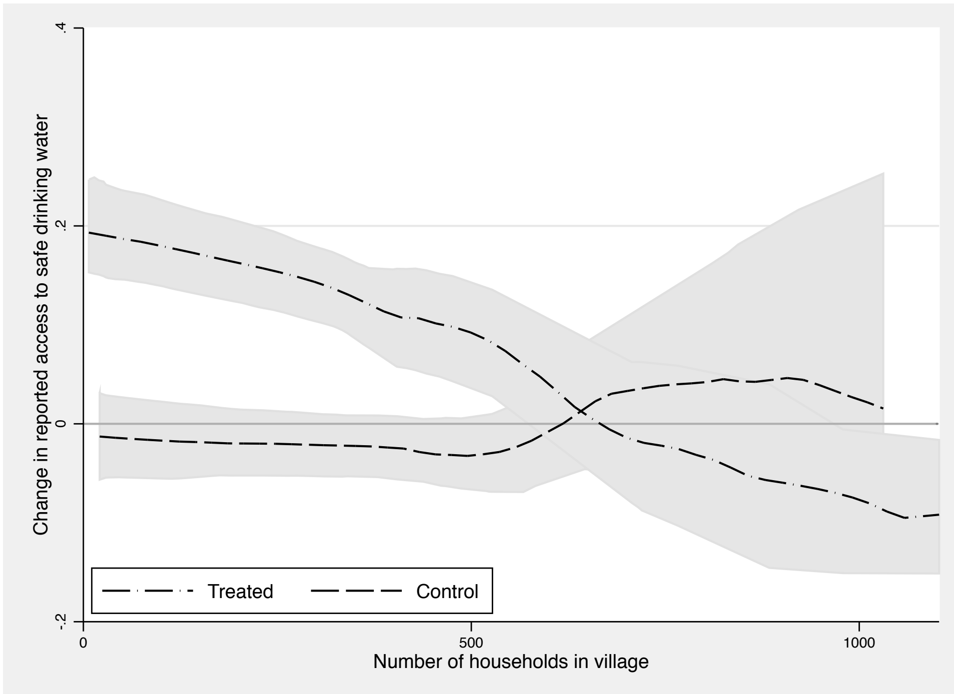
Stiglitz, J. (2002). Participation and development: Perspectives from the comprehensive development paradigm. *Review of Development Economics*, *6*(2), 163–182.

UNICEF. (2008, October). *Arsenic mitigation in bangladesh*.

World Bank. (2004). *World Development Report 2004: Making services work for poor people*. Washington, DC: The World Bank.

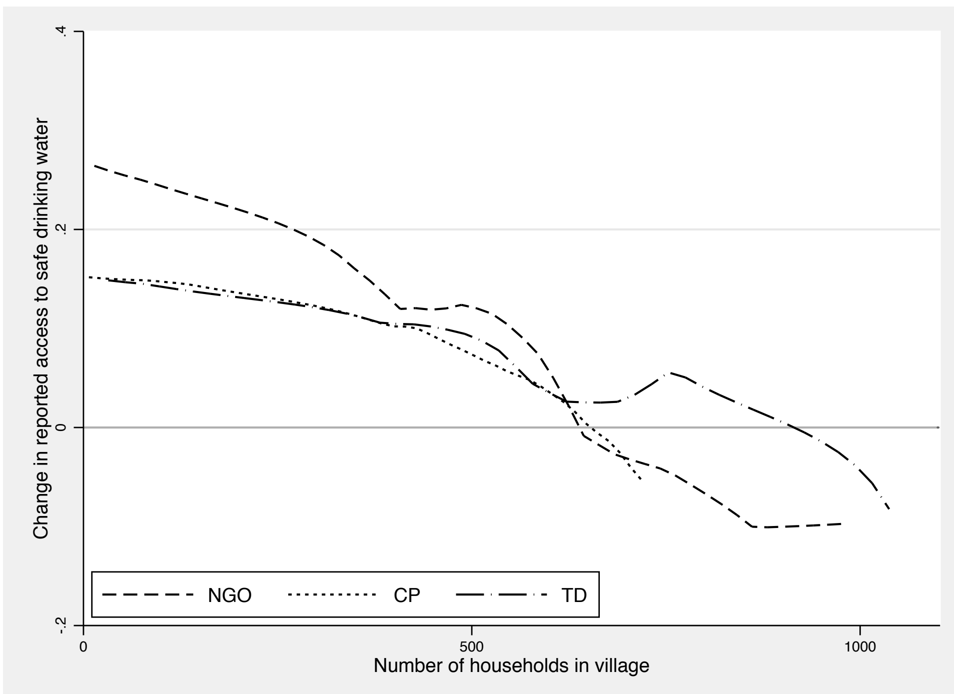
Figures and Tables

Figure 1: Heterogeneity in average treatment effect with village size



Graph shows results from a local linear regression of the change in access to safe drinking water on village size for treated and control villages. 90% confidence intervals are cluster bootstapped at the village level.

Figure 2: Heterogeneity in treatment effect by decision-making model with village size



Graph shows results from a local linear regression of the change in access to safe drinking water on village size for treated and control villages.

Table 1: Treated vs Control
Baseline Summary Statistics and Randomization Checks

		Control	Treated	Treatment - Control		
		(1)	(2)	(3)	(4)	(5)
Proportion of villages in Gopalganj	Mean	0.51	0.55			
	s.e.	(0.05)	(0.04)			
Proportion of villages in South Matlab	Mean	0.00	0.23***			
	s.e.	-	(0.04)			
No of households in village	Mean	245	221	-26	-29	-34
	s.e.	(21)	(17)	(29)	(27)	(31)
% of water sources arsenic contaminated	Mean	0.96	0.95	0.01	0.00	0.00
	s.e.	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reports using arsenic safe water	Mean	0.55	0.41***	-0.01	-0.03	-0.03
	s.e.	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
Changed source of drinking water due to arsenic in last 5 years?	Mean	0.49	0.35***	0.00	-0.02	-0.03
	s.e.	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Anyone in household has symptoms of arsenic poisoning?	Mean	0.009	0.009	-0.001	0.000	0.000
	s.e.	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)
Total value of household assets	Mean	570284	540165	-14573	-5071	-5865
	s.e.	(30362)	(21720)	(41233)	(37175)	(42881)
Access to electricity?	Mean	0.46	0.39	-0.05	-0.02	-0.03
	s.e.	(0.03)	(0.03)	(0.05)	(0.04)	(0.05)
Household head literate	Mean	0.61	0.60	0.01	0.00	0.00
	s.e.	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Household head Muslim	Mean	0.70	0.70	0.04	0.05	0.05
	s.e.	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)
Household head farmer	Mean	0.42	0.45	0.03	0.02	0.03
	s.e.	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Number of associations in community	Mean	6.25	6.30	-0.18	-0.18	-0.20
	s.e.	(0.14)	(0.15)	(0.22)	(0.19)	(0.22)
Number of collective actions in community	Mean	0.89	0.96	0.05	0.07	0.08
	s.e.	(0.08)	(0.09)	(0.05)	(0.06)	(0.06)
Number of villages		99	127	197	226	226
Number of households		3914	4976	7755	8890	8890

Note: P-values test significance of differences between treated and control villages, controlling for the different treatment proportions in Gopalganj, North and South Matlab in columns 3-5). Data in rows 1) and 2) come from project records. Data in rows 3) and 4) comes from data from the Bangladesh Arsenic Mitigation Water Supply Project. All other data is from baseline household surveys. Two villages are missing all baseline data. Standard errors (in parentheses) are robust, and clustered at the village level in rows 5) onwards.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Decision-making structures

Non-participatory	Top Down (TD)	Project staff took all project decisions, after an extended (typically 2-day) period of information gathering, using the following criteria to decide water source location: • public/convenient location • population density • existing safe water options
Participatory	Community Participation (CP)	The community took all project decisions using their own (unobserved) decision-making structures, following a community-wide information meeting led by project staff.
	NGO-Facilitated Community Participation (CP)	The community took all project decisions at a community- wide meeting, following smaller information meetings for different groups. We imposed two decision-making rules. If decisions made did not satisfy these rules, project staff did not implement the decisions: • Attendance at the community meeting had to include: at least 10 men, of which 5 had to qualify as poor; and at least 10 women, of which 5 had to qualify as poor. • Decisions had to be unanimous.

Table 3: Technologies to provide arsenic-safe drinking water

		Required community contribution per safe water source installed		
Technology	Cost	1	2	3
Deep tubewell (DTW)	50000	4500	6000	7500
Shallow tubewell (STW)	20000	3000	3500	4000
Arsenic-Iron Removal Plant (AIRP)	60000	6000	7500	N/A
Deep-set tubewell (DSTW)	60000	6000	7500	N/A

Note: All prices in Bangladeshi Taka. 1 US$≈ 80BDT.

Table 4: Assignment to decision-making structure
Baseline Summary Statistics and Randomization Checks

		NGO (1)	CP (2)	TD (3)
Proportion of villages in Gopalganj	Mean	0.55	0.55	0.56
	s.e.	(0.08)	(0.08)	(0.08)
Proportion of villages in South Matlab	Mean	0.24	0.21	0.23
	s.e.	(0.07)	(0.06)	(0.06)
No of households in village	Mean	213	213	236
	s.e.	(33)	(24)	(32)
% of water sources arsenic contaminated	Mean	0.96	0.95	0.95
	s.e.	(0.01)	(0.01)	(0.01)
AIRPs only feasible technology	Mean	0.10	0.14	0.14
	s.e.	(0.05)	(0.05)	(0.05)
Reports using arsenic safe water	Mean	0.36	0.42	0.44
	s.e.	(0.05)	(0.05)	(0.05)
Changed source of drinking water due to arsenic in last 5 years?	Mean	0.32	0.35	0.37
	s.e.	(0.05)	(0.05)	(0.05)
Anyone in household has symptoms of arsenic poisoning?	Mean	0.004**	0.009	0.012*
	s.e.	(0.002)	(0.003)	(0.003)
Total value of household assets	Mean	543364	548174	529180
	s.e.	(39597)	(42208)	(30413)
Access to electricity?	Mean	0.37	0.39	0.42
	s.e.	(0.05)	(0.05)	(0.05)
Household head literate	Mean	0.60	0.58	0.62
	s.e.	(0.03)	(0.03)	(0.02)
Household head Muslim	Mean	0.68	0.70	0.73
	s.e.	(0.07)	(0.06)	(0.06)
Household head farmer	Mean	0.44	0.46	0.44
	s.e.	(0.03)	(0.02)	(0.02)
Number of associations in community	Mean	6.36	6.05	6.47
	s.e.	(0.25)	(0.18)	(0.31)
Number of collective actions in community	Mean	0.91	1.00	0.98
	s.e.	(0.14)	(0.16)	(0.15)
	Number of villages	42	42	43
	Number of households	1638	1635	1703

Note: P-values test significance of the difference between model and other treated villages. Data from household surveys except rows 1), 2) and 5) which come from project records and rows 3) and 4) which come from the Bangladesh Arsenic Mitigation Water Supply Project. Baseline data for one CP village is missing. Standard errors (in parentheses) are robust, and clustered at the village level in rows 5) onwards. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Participation in project decision-making

		No. of participants	Fraction of meeting participants: Female	Low s.e. status	No 2° education
		(1)	(2)	(3)	(4)
All treated	Mean	30.7	0.27	0.42	0.60
	s.e.	(1.1)	(0.02)	(0.02)	(0.02)
NGO	Mean	33.6	0.31	0.44	0.64
	s.e.	(2.5)	(0.03)	(0.04)	(0.03)
CP	Mean	29.9	0.29	0.46	0.55
	s.e.	(1.4)	(0.03)	(0.02)	(0.03)
TD	Mean	28.5	0.21	0.36	0.61
	s.e.	(1.8)	(0.03)	(0.03)	(0.03)
NGO = CP	p-value	0.200	0.615	0.570	0.031
CP = TD	p-value	0.561	0.073	0.006	0.154
TD = NGO	p-value	0.111	0.023	0.085	0.490
	N	126	123	119	122

Note: P-values test pairwise significance of the difference between the means across models indicated, from a regression of the outcome variable on indicators for the three types of treatment (with no constant). Robust standard errors shown in parentheses. In two villages, meetings were not held.

Table 6: Project Outcomes

		Outcome Variable		
		Water sources installed	Installed in public places	Number of contributors
		(1)	(2)	(3)
Panel A: All villages				
All treated	Mean	2.14	1.38	7.81
	s.e.	(0.10)	(0.10)	(0.83)
NGO	Mean	2.21	1.19	9.37
	s.e.	(0.16)	(0.14)	(1.62)
CP	Mean	2.21	1.00	5.40
	s.e.	(0.17)	(0.14)	(0.93)
TD	Mean	2.00	1.93	8.67
	s.e.	(0.18)	(0.18)	(1.63)
NGO = CP	p-value	1.000	0.334	0.036
CP = TD	p-value	0.394	0.000	0.084
TD = NGO	p-value	0.376	0.002	0.764
	N	127	127	126
Panel B: Villages where tubewells feasible				
All treated	Mean	2.45	1.58	9.02
	s.e.	(0.08)	(0.10)	(0.92)
NGO	Mean	2.46	1.30	10.61
	s.e.	(0.13)	(0.14)	(1.74)
CP	Mean	2.53	1.14	6.11
	s.e.	(0.14)	(0.15)	(1.03)
TD	Mean	2.36	2.31	10.33
	s.e.	(0.16)	(0.15)	(1.82)
NGO = CP	p-value	0.718	0.440	0.028
CP = TD	p-value	0.428	0.000	0.046
TD = NGO	p-value	0.624	0.000	0.912
	N	109	109	108

Note: P-values test pairwise significance of the difference between the means across models indicated, from a regression of the outcome variable on indicators for the three types of treatment (with no constant). Robust tandard errors shown in parentheses. In villages where no water sources were installed, the number of contributors is coded as zero. In one village, the number of contributors was not recorded.

Table 7: Estimates of average treatment effect

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel A: Reported access to safe drinking water at baseline								
Treated	Coefficient	-0.141***	-0.011	-0.031	-0.181***	-0.017	-0.042	-0.024
	s.e.	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.10)
Control	Coefficient	0.547	0.830	0.837	0.609	0.833	0.842	0.253
	s.e.	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.07)
	First-stage F-test			1340			896	
	N	8695	7608	8695	7375	6288	7375	1242
Panel B: Reported access to safe drinking water at follow-up								
Treated	Coefficient	0.022	0.141***	0.132***	0.001	0.148***	0.137***	-0.007
	s.e.	(0.05)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.08)
Control	Coefficient	0.533	0.820	0.823	0.607	0.817	0.821	0.201
	s.e.	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.06)
	First-stage F-test			1335			896	
	N	8442	7387	8442	7168	6113	7168	1201
Panel C: Change in reported access to safe drinking water								
Treated	Coefficient	0.164***	0.148***	0.157***	0.182***	0.164***	0.175***	0.004
	s.e.	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.09)
Control	Coefficient	-0.014	-0.012	-0.015	-0.003	-0.017	-0.021	-0.042
	s.e.	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
	First-stage F-test			1337			897	
	N	8427	7375	8427	7154	6102	7154	1200
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Treatment is instrumented using synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 8: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel A: Reported access to safe drinking water at baseline								
NGO	Coefficient	-0.19***	-0.06	-0.07	-0.24***	-0.08	-0.09	-0.10
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.09)
CP	Coefficient	-0.13**	0.01	-0.02	-0.17**	0.00	-0.04	0.04
	s.e.	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)	(0.06)	(0.17)
TD	Coefficient	-0.10	0.02	0.00	-0.13*	0.03	0.01	-0.04
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.10)
Constant	Coefficient	0.55	0.83	0.84	0.61	0.83	0.84	0.25
	s.e.	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.07)
	NGO = CP	0.437	0.295	0.460	0.395	0.257	0.443	0.389
	CP = TD	0.708	0.916	0.673	0.563	0.683	0.456	0.609
	TD = NGO	0.239	0.199	0.208	0.151	0.106	0.119	0.540
	NGO = pooled	0.258	0.180	0.252	0.185	0.114	0.180	0.327
	CP = pooled	0.819	0.587	0.851	0.878	0.673	0.988	0.480
	TD = pooled	0.372	0.423	0.329	0.246	0.241	0.183	0.873
	N	8695	7608	8695	7375	6288	7375	1242
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 8, continued: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel B: Reported access to safe drinking water at follow-up								
NGO	Coefficient	0.03	0.15***	0.16***	0.01	0.16***	0.16***	-0.13*
	s.e.	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
CP	Coefficient	0.00	0.14***	0.11**	-0.03	0.14**	0.10*	0.09
	s.e.	(0.06)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.13)
TD	Coefficient	0.03	0.13***	0.13***	0.02	0.15***	0.15***	-0.02
	s.e.	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.09)
Constant	Coefficient	0.53	0.82	0.82	0.61	0.82	0.82	0.20
	s.e.	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.06)
	NGO = CP	0.702	0.840	0.430	0.657	0.725	0.308	0.091
	CP = TD	0.697	0.860	0.634	0.499	0.875	0.372	0.473
	TD = NGO	0.998	0.708	0.699	0.809	0.823	0.811	0.178
	NGO = pooled	0.823	0.748	0.504	0.893	0.745	0.467	0.042
	CP = pooled	0.660	0.978	0.461	0.528	0.764	0.272	0.217
	TD = pooled	0.817	0.740	0.968	0.584	0.963	0.704	0.958
	N	8442	7387	8442	7168	6113	7168	1201
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 8, continued: Estimates of treatment effect by decision-making model

		OLS (1)	OLS (2)	IV (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)
Panel C: Change in reported access to safe drinking water								
NGO	Coefficient	0.22***	0.21***	0.22***	0.25***	0.24***	0.25***	-0.05
	s.e.	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.09)
CP	Coefficient	0.14***	0.13***	0.13***	0.15***	0.13***	0.14***	0.03
	s.e.	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.15)
TD	Coefficient	0.13***	0.11**	0.13***	0.15***	0.12**	0.14***	0.02
	s.e.	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.12)
Constant	Coefficient	-0.01	-0.01	-0.02	0.00	-0.02	-0.02	-0.04
	s.e.	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.05)
	NGO = CP	0.178	0.238	0.149	0.116	0.143	0.083	0.653
	CP = TD	0.946	0.783	0.969	0.954	0.765	0.930	0.947
	TD = NGO	0.151	0.160	0.155	0.109	0.104	0.106	0.636
	NGO = pooled	0.119	0.150	0.110	0.078	0.088	0.064	0.568
	CP = pooled	0.429	0.556	0.358	0.338	0.432	0.250	0.786
	TD = pooled	0.356	0.302	0.388	0.298	0.231	0.329	0.850
	N	8427	7375	8427	7154	6102	7154	1200
	Feasible technology	All	All	All	Tubewell	Tubewell	Tubewell	AIRP
	Controls for upazila	No	Yes	Yes	No	Yes	Yes	-
	Includes S. Matlab	Yes	No	Yes	Yes	No	Yes	-
	Control villages	All	All	All	Matched	Matched	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab in columns 3) and 6). In columns 4) to 7) the control group is matched to the subset of treated villages using baseline propensity score matching. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Standard errors (in parentheses) are robust and clustered at the village level.
*** p<0.01, ** p<0.05, * p<0.1.

Table 9: Robustness checks

		OLS (1)	OLS (2)	IV (3)	IV (4)	IV (5)	OLS (6)	IV (7)	IV (8)	IV (9)	IV (10)
NGO	Coefficient	0.24***	0.37***	0.10	0.25***	0.27***	0.24***	0.46***	0.28***	0.28***	0.30***
	s.e.	(0.05)	(0.09)	(0.07)	(0.06)	(0.06)	(0.06)	(0.13)	(0.07)	(0.07)	(0.07)
CP	Coefficient	0.15***	0.19***	0.08	0.14***	0.15***	0.14***	0.18	0.15***	0.16**	0.13***
	s.e.	(0.04)	(0.07)	(0.07)	(0.05)	(0.05)	(0.04)	(0.16)	(0.06)	(0.07)	(0.05)
TD	Coefficient	0.14***	0.21***	0.06	0.14***	0.16***	0.12**	0.21*	0.16***	0.20***	0.14***
	s.e.	(0.04)	(0.07)	(0.06)	(0.05)	(0.05)	(0.05)	(0.12)	(0.06)	(0.07)	(0.05)
Constant	Coefficient		-0.04	0.01	-0.02	-0.03		-0.03	-0.03	0.20	0.02
	s.e.		(0.04)	(0.02)	(0.02)	(0.02)		(0.06)	(0.02)	(0.04)	(0.03)
	NGO = CP	0.118	0.055	0.784	0.083	0.084	0.165	0.133	0.098	0.117	0.020
	NGO = TD	0.111	0.107	0.637	0.106	0.108	0.122	0.117	0.151	0.280	0.027
	NGO = pooled	0.080	0.052	0.674	0.064	0.065	0.106	0.077	0.086	0.130	0.013
	N	4017	3201	3953	7342	7784	4660	8347	8357	7154	7154
	Measure of access	Reported	Reported	Reported	Reported	Reported	Reported	Reported	Reported	Verified	Combined
	Feasible technology	Tubewell	Tubewell	Tubewell	Tubewell	Tubewell	All	All	All	Tubewell	Tubewell
	Sample	All	Gopalganj	Matlab	All	All	All	All	All	All	All
	Upazila Controls	No	-	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
	Additional controls	No	No	No	No	No	Technology	Vill. size	Assets	No	No
	Control villages	None	Matched	All	Alternative	All	None	All	All	Matched	Matched

Note: Model is instrumented using model interacted with synthetic assignment to treatment in South Matlab where indicated. Survey weights are applied so that each village counts equally in the analysis. Controls for upazila include an indicator for Gopalganj, and an indicator for South Matlab, where included. Where additional controls are included, the interactions between the controls and the treatment indicators are also
Standard errors (in parentheses) are robust and clustered at the village level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendices

Appendix A: Data

Variable Construction

We asked the households to list all the sources of water they used for drinking and cooking. In the analysis, we focus on the most important source of water for drinking and cooking, which we asked households to list first. We also asked households to report the percentage of water for drinking and cooking that they obtained from each source, but results based on the source from which households report drawing the largest percentage of water are unstable between baseline and followup, whereas the results based on the first-listed water source are more consistent. This may be attributable to slight differences in the way in which the question was asked as to whether the question referred to water used for drinking only or drinking and cooking.

Reports using safe drinking water If the household reports using a tubewell, we code the household as reporting using safe water if they report that the source is arsenic-safe, and reporting unsafe water if it is unsafe or if they dont know the source's safety. If the household reports using an unsafe source with respect to bacterial contamination (i.e. a dug well or surface water), we code the household as reporting using unsafe water. Some sources can be presumed to be safe from both bacterial and arsenic contamination (e.g. AIRPs, PSF, rainwater, deep-set tubewells). In these cases, we code the household as reporting using safe water unless the household reports that the water is unsafe. The numbers of households using these sources is small. If the household reports using any other source, we code the household as reporting using safe water if they report that the source is safe, and reporting unsafe water if it is unsafe or if they dont know the source's safety status.

Reports using verified safe drinking water Many tubewells in Bangladesh have been tested for arsenic safety in the past and marked with green (safe) or red (unsafe). If households reported using a tubewell, enumerators visited the tubewell to confirm whether it was marked safe, unsafe, or not marked, as long as the tubewell was less than 5 minutes walk away. However, an early version of the baseline survey used did not include this question, so this information is missing for some villages at baseline. We code this variable as follows. We always code households reporting an unsafe source as not using a verified source of safe drinking water. If the household reports using a source that can otherwise be presumed to be safe (i.e. rainwater), we code the household as using a verified source of safe drinking water, unless they report it to be unsafe. If they report using a tubewell, we code the house as using a verified source of safe drinking water if the source is verifiably safe, and not using a verified source of safe drinking water if the tubewell is either marked unsafe or can't be verified safe (either because it is too far, or because it is unmarked, or because the question wasn't asked in this village at baseline). If the household reports using any other source, we code the household as reporting using a verified source safe water if the enumerators recorded that the source could be verified safe, and reporting using an unverified or unsafe source if the source is unsafe or if it cannot be verified.

Combined measure We combine these measures by using the verified data, when the safety of the source could be verified, and the reported data, when the safety of the source could not be verified.

Construction of matched control groups in Gopalganj

Tubewells are not feasible where there is a rocky layer separating the surface from the arsenic-safe deep aquifer, which cannot be penetrated with local drilling technologies. This was only a problem in this study in Gopalgnaj, as tubewells of varying kinds were feasible in all villages in Matlab. The following discussion is therefore limited to Gopalganj. There is substantial spatial correlation in the location of these rocky layers.

In Gopalganj, there was no overall problem with random assignment to treatment. Appendix Table B1, columns 1) and 2) confirm this; of 12 tests comparing treated to control villages in Gopalganj, only one shows statistically significant differences at the 10% level, which is approximately what we would expect due to chance. However, when we compare either the AIRP villages (column 3) or the tubewell villages (column 6), to the full group of control villages, there is some evidence that feasible technology is correlated with other village level characteristics. In column 3), only 1 of 12 tests shows statistically significant differences at the 10% level, but in column 6), 1 test shows statistically significant differences at the 5% level, and additionally another shows statistically significant differences at the 10% level. To be conservatve, we construct matched control groups for the AIRP villages and tubewell villages, primarily exploiting the spatial correlation in location of the rocky layer.

In Gopalganj, there are 18 unions (the smallest rural administrative and local government units in Bangladesh). In three of these unions, only AIRPs were feasible in all treated villages. In 4 unions, tubewells were feasible in all treated villages. First, we assign the five control villages in unions where only AIRPs were feasible to the AIRP-matched control group. Second, we assign the 18 control villages in unions where tubewells were always feasible to the tubewell-matched control group. For the remainder of the unions — for which tubewells were feasible in a fraction of the villages — we use two strategies to identify matched control villages.

First, we construct a propensity score index by running a logit regression of an indicator for whether tubewells (or only AIRPs) were feasible on a set of baseline characteristics which we observe for both treated and control groups and which are correlated feasible technology in the treatment group. These baseline characteristics are: union fixed effects, access to electricity, number of tubewells operative, number of tubewells arsenic contaminated, reported access to safe drinking water at baseline, verifiability of arsenic safety of primary reported source, verified access to safe drinking water at baseline, fraction of village changing source of safe drinking water in the preceding five years due to arsenic contamination; whether any member of the household has arsenic poisoning; mean literacy of household heads; and number of coordinated actions in the village. We then assign the remaining villages to the relevant matched control group where their propensity score is greater than 0.5 (although in reality the propensity scores are strongly clustered around 0 and 1). This process assigns a total of 16 villages to the AIRP-matched control group, and 33 villages to the tubewell-matched control group. 3 villages are not assigned to either control group, and 2 villages are assigned to both groups. Columns 4) and 7) show the characteristics of the groups constructed in this way. No tests show statistically significant differences between these groups and the AIRP or tubewell villages from the treated groups.

The second method we use is to assign the villages in each union to the tubewell-matched and AIRP-matched control groups according to the observed, union-level probability of AIRP and tubewell feasibility. We do this by generating a random number for each control village and assigning them to the relevant control group if the random number is less than the proportion of treated villages in that union in which the technology was feasible. This process assigns 9 villages to the AIRP-matched control villages and 40 villages to the tubewell-matched control villages. One village is not assigned to either control group, and by construction none are assigned to both groups. Columns 5) and 8) show the characteristics of the groups constructed in this way. 3 tests show statistically significant differences between the AIRP-matched control villages and the AIRP villages from the treated groups, but no tests show statistically significant differences between the tubewell-matched control villages constructed in this way and the tubewell villages from the treated groups.

For the main specification, we therefore use the matched control group created using propensity score matching, but we will compare these results to results using the full control group and the second matching strategy in robustness checks.

Appendix B: Figures and Tables

Table B1: Baseline Summary Statistics for AIRP and non-AIRP villages in Gopalganj

	Treated	Control	AIRP	AIRP Control 1	AIRP Control 2	Tubewell	Tubewell Control 1	Tubewell Control 2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No of households in village	262	257	338	268	271	242	255	263
	(26)	(33)	(71)	(44)	(57)	(27)	(45)	(41)
% of water sources arsenic contaminated	0.97	0.96	0.97	0.97	0.99*	0.96	0.96	0.96
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reports using arsenic safe water	0.27	0.26	0.23	0.25	0.26	0.27	0.27	0.26
	(0.03)	(0.04)	(0.07)	(0.07)	(0.11)	(0.04)	(0.04)	(0.04)
Changed source of drinking water due to arsenic in last 5 years?	0.21	0.20	0.22	0.21	0.19*	0.19	0.18	0.19
	(0.03)	(0.03)	(0.06)	(0.07)	(0.10)	(0.03)	(0.04)	(0.04)
Anyone in household has symptoms of arsenic poisoning?	0.006	0.009	0.005	0.008	0.006	0.006	0.003	0.008
	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)
Total value of household assets	469230	512028	568106	531573	527437	436844	493045	514595
	(21450)	(40937)	(47317)	(82727)	(146437)	(23279)	(47148)	(41289)
Access to electricity?	0.37	0.45	0.60*	0.68	0.63	0.31**	0.35	0.41
	(0.04)	(0.04)	(0.07)	(0.03)	(0.05)	(0.04)	(0.05)	(0.05)
Household head literate	0.58	0.56	0.54	0.49	0.45	0.59	0.62	0.58
	(0.03)	(0.04)	(0.06)	(0.07)	(0.08)	(0.03)	(0.04)	(0.04)
Household head Muslim	0.55	0.48	0.66	0.46	0.52	0.52	0.46	0.48
	(0.05)	(0.06)	(0.11)	(0.10)	(0.14)	(0.06)	(0.08)	(0.07)
Household head farmer	0.50	0.46	0.45	0.39	0.39	0.51*	0.49	0.47
	(0.02)	(0.02)	(0.04)	(0.03)	(0.05)	(0.02)	(0.03)	(0.02)
Number of associations in community	6.76	6.86	6.65	7.18	7.62	6.77	6.79	6.82
	(0.23)	(0.22)	(0.36)	(0.50)	(0.70)	(0.29)	(0.23)	(0.22)
Number of collective actions in community	0.23	0.14*	0.26	0.12	0.10*	0.22	0.15	0.14
	(0.04)	(0.02)	(0.08)	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)
Number of villages	70	50	16	16	9	52	33	38
Number of households	3859	1969	1102	627	347	2677	1304	1504

Note: Stars indicate significance of tests of difference of means: between treated and control villages in Gopalganj in column 2); between AIRP and all control villages in column 3); between AIRP and matched control villages in columns 4) and 5); between tubewell and all control villages in column 6; and between tubewell and matched control villages in columns 7) and 8). Data in rows 1) and 2) come from the Bangladesh Arsenic Mitigation Water Supply Project. The remaining data are from household surveys. Standard errors (shown in parentheses) are clustered at the village level. *** p<0.01, ** p<0.05, * p<0.1.