Improving state effectiveness in environmental risk mitigation: Pre-Analysis Plan

Prabhat Barnwal[1], Michael Carlos Best[2], Md. Ahasan Habib[3], Seung Min Kim[4], Alexander van Geen[5], and Jack Willis[6]

[1]Department of Economics, Michigan State University
[2]Department of Economics, Columbia University
[3]NGO Forum for Public Health
[4]School of International and Public Affairs, Columbia University
[5]Lamont-Doherty Earth Observatory, Columbia University
[6]Department of Economics, Sciences Po

September 4, 2026

This text represents the intentions of the research team at the time it is filed. We may deviate from this plan if unexpected issues arise, but we will report a "populated PAP" if we materially deviate from it (Banerjee et al., 2020).

Abstract

This study aims to understand the barriers to improving the targeting of policy instruments. Focusing on the case of arsenic-safe governmental tubewells in rural Bangladesh, we conducted a lab-in-the-field experiment to evaluate the separate and joint impacts of two obstacles: lack of information and incentive misalignment. This pre-analysis plan presents the details of our experimental design and survey and our plan for how to analyze the resulting data.

Contents

1 Introduction

In many developing countries, effectively targeting policy instruments remains challenging due to the difficulty of identifying those most in need (Alatas et al., 2012; Blumenstock et al., 2015; Haushofer et al., 2025). Yet, even when such information barrier is removed and policymakers can perfectly identify the most impactful resource allocation, corruption and misaligned incentives often prevent the realization of such optimal allocation in practice (Olken and Pande, 2012).

This paper explores the separate and joint effects of information and incentive misalignment in the optimal targeting of policy. We focus on rural Bangladesh, where more than 60 million people are exposed to drinking water with arsenic levels higher than the WHO guideline–and where the suboptimal allocation of public, arsenic-safe governmental wells have been subject to significant elite capture (Human Rights Watch, 2016; Mobarak and van Geen, 2019). We combine a quasi nation-wide arsenic field test data comprising 6.3 million wells with machine learning models of well-water arsenic to develop an interactive website displaying localized arsenic contamination status ("arsenic information dashboard"). We then conduct a lab-in-the-field experiment involving 200 local policymakers who hold discretionary authority in allocating the governmental tubewells. The policymakers are randomly assigned into three treatments: (1) an "information treatment" involving the arsenic information dashboard, (2) an "incentive" treatment, where the policymakers are incentivized to prioritize high-arsenic neighborhoods, and (3) a joint intervention combining the two. The experiment is designed to be highly realistic and high-stakes, using real applications from villagers for public tubewells and involving a real probability of the representatives' stated preferences being realized.

This pre-analysis plan describes the set-up and the structure of the lab-in-the-field experiment in detail, and presents the survey that accompanied the experiments. We also discuss–and commit to–our plans for how to analyze the resulting data that will be collected from the experiment.

2 Well-water Arsenic in Rural Bangladesh

2.1 Arsenic in Rural Bangladesh: An overview

Naturally occurring arsenic in shallow groundwater is among the largest environmental health risks in the developing world, affecting hundreds of millions of people globally (Podgorski and Berg, 2020). In rural Bangladesh, following a decades-long public health campaign that moved rural households away from microbially contaminated surface water sources and towards groundwater drawn from shallow hand-pumped tube wells, an unanticipated consequence emerged: a large share of the shallow Holocene aquifers underlying the Bengal Basin is naturally enriched in dissolved arsenic (British Geological Survey and Department of Public Health Engineering, 2001). More than 60 million people in Bangladesh drink water with arsenic concentrations above the World Health Organization guideline of 10 μg/L (Flanagan et al., 2012), and exposure of this magnitude is estimated to generate on the order of 100,000 deaths annually (Quansah et al., 2015).

The health and economic consequences of chronic exposure are well documented. Long-run cohort evidence from Bangladesh finds that individuals drinking water with arsenic concentrations above 100 μg/L face roughly a 100% increase in the all-cause mortality risk of unexposed individuals (Argos et al., 2010). Beyond mortality, chronic arsenicosis is associated with adverse pregnancy and infant health outcomes (Quansah et al., 2015) and with substantial losses in labor productivity (Pitt et al., 2021). Further, recent evidence (Wu et al., 2025) indicates that switching away from a contaminated well measurably lowers mortality risk, highlighting the importance of risk mitigation measures.

Two features of the geography of well-water arsenic contamination are central to the design of this study.

Depth. Arsenic concentrations are strongly related to sand characteristics of sediments, and hence to tube well depth. Shallow tube wells accessing Holocene era aquifers–the overwhelming majority of privately installed wells–draw from the aquifers most likely to be contaminated, whereas deeper aquifers from the Pleistocene era sediments are typically arsenic-safe (Horneman et al., 2004).

In practice, there are two ways through which a villager could access arsenic-safe aquifers. First, in cases where the safe aquifer exists at depths accessible to private tube well installers (with inexpensive drilling technology), a household can opt to install a well to sufficiently deep levels upon new installation. Second, when the safe aquifer exists beyond the reach of such private, village-level installers (i.e., beyond 300 feet), then a governmental well (called "DPHE wells," from the Department of Public Health Engineering (DPHE), the governmental body in charge of drinking water source policy) is necessary. Nevertheless, a DPHE well costs roughly USD 1,000 per well, an order of magnitude larger than what a typical rural household will finance privately. This makes publicly-installed wells a coveted resource subject to heavy oversubscription. The allocation of this DPHE wells is the target of our intervention.

Fine-grained spatial heterogeneity. Contamination varies sharply not only across districts and upazilas but *within* unions, across neighboring *mauzas* (collection of one or more villages), and across neighboring villages. This spatial granularity is economically relevant for allocation. Because the number of wells a union receives is fixed by budgetary convention (Section 2.2), the public health return to a given year's allocation is determined almost entirely by *which villages within the union* receive the wells. Two allocations of identical size within the same union can differ significantly in the number of arsenic-exposed households they reach.

Until recently, the fine-grained information required to act on this heterogeneity did not exist in usable form. That changed with DPHE's Arsenic Risk Reduction Program (ARRP), a quasi-nationwide field-testing campaign that tested approximately 6.3 million drinking-water wells across roughly 300 upazilas between 2021 and 2023, recording each well's geographic coordinates, depth, ownership type (public or private), and field-kit arsenic reading.[1] ARRP makes village-level contamination rates measurable at national scale for the first time.

ARRP is nonetheless not a census of drinking-water wells. Testing was capped at 2,500 wells per union (Figure 1), with the consequence that roughly one third of villages nationwide contain no tested well. Since the allocation problem we study is precisely the within-union, across-village one, these gaps are not innocuous: a decision aid built on tested villages alone would systematically omit a third of the candidate sites. We therefore gap-fill untested villages using a machine-learning predictor of well-level arsenic safety — a random forest (Breiman, 2001) with 500 trees, taking well longitude, latitude, and depth as inputs — fit separately by district using wells from the district and its neighbors so as to preserve spatial continuity across district boundaries.[2] Model performance is evaluated by *grouped* cross-validation: rather than splitting individual wells, we hold out all wells in a random third of villages, aggregate predictions to the village level, and compare predicted to realized shares of unsafe wells. This mirrors the spatial cross-validation practice in the ecological prediction literature (Meyer et al., 2019) and rules out the possibility that measured accuracy reflects memorization of village identities rather than genuine out-of-sample spatial prediction. Across nearly all districts, out-of-sample RMSE for village-level contamination rates is smaller for the random forest than for a kriging-type OLS benchmark using polynomial functions of coordinates and depth, and is smaller on average

[1] Field-kit readings were recalibrated against laboratory analysis of a random subset of samples processed in DPHE laboratories; see Khan et al. (2025). After cleaning to remove duplicates, coordinate outliers, and non-drinking-water wells, we work with approximately 6.0 million wells.

[2] The outcome is a graded safety score constructed from the field-kit thresholds, taking the value 1 for readings at or below 10 μg/L, 0.5 for readings of 25 or 50 μg/L, and 0 for readings strictly above 50 μg/L, following the laboratory validation in Khan et al. (2025).

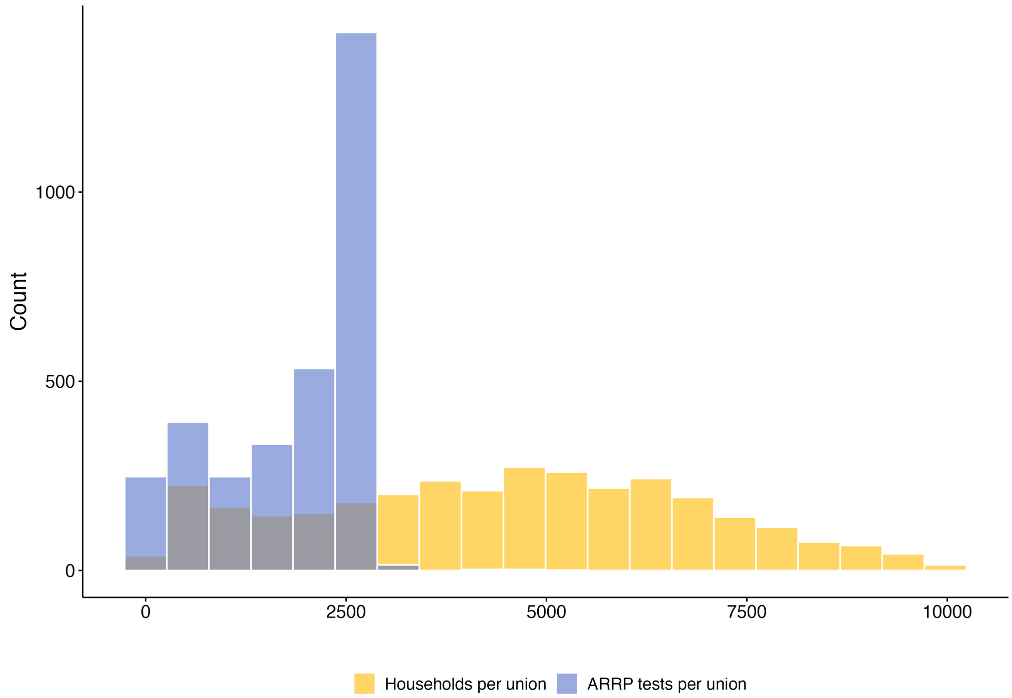


Figure 1. Incomplete arsenic testing in the ARRP campaign

Notes: Figure illustrates the incomplete coverage of wells in the Arsenic Risk Reduction Program (ARRP) field arsenic test campaign. The distribution of well tests per union are shown in blue, versus the number of households of the same unions in yellow.

than the standard deviation of village-level contamination rates — i.e. the predictions are informative relative to the unconditional mean.

2.2 Current policy approaches to arsenic risk mitigation

The Government of Bangladesh has recognized arsenic contamination as a public health priority for over two decades and has committed substantial resources to mitigation. The operational responsibility rests with the DPHE, the government body charged with allocating and installing arsenic-safe public wells. The core instrument is the publicly financed deep tube well, installed free of charge to the recipient community.

The allocation process. Allocation is application-driven and highly decentralized. In each fiscal year, villagers submit applications to their local union parishad (UP) office and/or the local DPHE office requesting installation of a well in their village. Bangladesh is divided into roughly 500 *upazilas*, each comprising on average about 70,000 households; each upazila is subdivided into *unions* of roughly 5,000 households, and each union con-

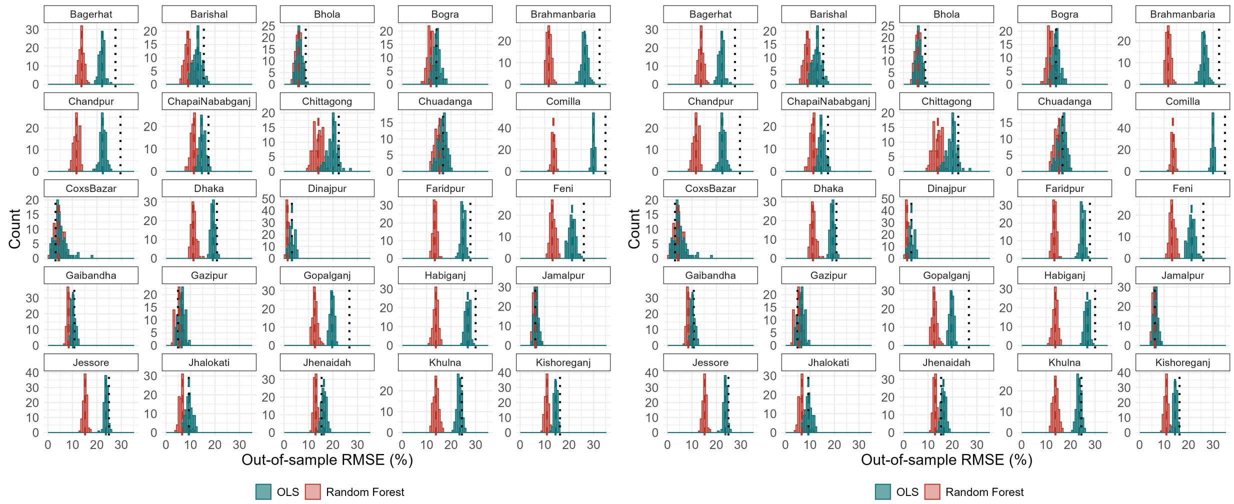


Figure 2. Out-of-sample performance of machine learning-based downscaling approach

Notes: Figure compares, in all districts with ARRP data, the out-of-sample RMSE (root mean squared error) of a krigging OLS model (green) versus that of a random forest-based downscaling model (red). The histogram is drawn by 50 different random sampling of sample draws, with sample drawn at the village level to avoid underestimations of the true bias rate. The standard deviation of village-level arsenic contamination shown in dotted black vertical line.

tains on average 13 villages. Within each union, a single Union Parishad chairperson (henceforth "union leader"), the elected representative at the union level,[3] decides which of the submitted applications applications will receive half of the budgeted DPHE wells. The remaining half of the budgeted wells are decided by the local Member of Parliament (MP), who represents on the order of 50,000–70,000 households. While the union leaders and MPs are formally advised by the upazila-level DPHE engineer, the engineers' roles have traditionally been relegated to managing the logistics of the wells' installation and their subsequent maintenance.[4]

The equity norm and the locus of discretion. Historically, the upazila-wide well budget has been divided evenly across the upazila's unions. This is a strong and durable equity convention: in past years each union within an upazila has received the same number of

[3]Note that after the July Uprising of 2024, many union leaders–who are in principle apolitical–went into hiding, for their suspected ties to the deposed Awami League party. Under such cases, the elected union leaders were substituted by an interim chairperson who were either assigned from a district-level bureaucratic, nominated among the union parishad's village-level representatives, or were taken over by the UP secretary, a bureaucrat formerly charged with day-to-day operations of the union.

[4]Upazila DPHE offices are formally headed by Assistant Engineers (AEs), the entry-level cadre position for engineers who have passed the Bangladesh Civil Service examination. In practice most upazila offices are staffed by Sub-Assistant Engineers (SAEs), recruited through DPHE-specific examinations into non-cadre positions; the upazila-level officer is thus referred to colloquially as the SAE. We use "upazila engineer" throughout.

wells, and we do not expect it to be violated over the study period.[5] The consequence is analytically convenient and substantively important: because the extensive margin (how many wells a union receives) is fixed by convention, essentially all of the discretion that determines the public health return to the annual allocation operates on the intensive margin — the choice of *which mauzas (villages) within the union* receive that union's wells. Union leaders and the upazila engineer face a within-union targeting problem, and that is the decision this study intervenes on. In the fiscal year covering our lab-in-the-field experiment, DPHE policy sets the union-level budget at four wells.

Two barriers to efficient allocation. Despite the scale of public investment, the realized allocation of DPHE deep wells falls well short of the total reduction in arsenic exposure the same budget could achieve. Administrative records and prior research document that placement bears little relation to measured contamination and is vulnerable to elite capture (van Geen et al., 2016; Mobarak and van Geen, 2019). We take two distinct barriers to be operative, and the experiment is designed to separate them.

First, an *information* barrier. Even where decision-makers are motivated to target high-arsenic villages, they may not know which villages these are. Although the ARRP campaign has made within-union contamination patterns measurable, that information has not been translated into a form usable by the officials who actually allocate wells.

To assess the existence of policy-relevant arsenic information at the union level, we elicit predictions of *mauza*-level arsenic contamination rate from DPHE upazila-level engineers, who nominally advise the union leaders on arsenic levels and local water quality in general. A comparison of the engineers' predictions with true arsenic contamination levels show that, however, while engineers exhibit acceptable knowledge of which *unions* in their upazila are high in arsenic, but pronounced deficiencies in identifying high-arsenic *villages within* a union (Appendix Figure B3) — precisely the margin on which their discretion operates. Engineers asked to estimate mauza-level contamination rates (for ARRP-tested mauzas) produced estimates far from the measured values (Appendix Figure B3). This strongly suggests that the union leaders do not possess the relevant information in prioritizing high-arsenic mauzas within their unions.

Second, an *incentives* barrier. The upazila engineer is an advisor, not a decision-maker, and formal authority over which applications are approved rests with elected representatives (i.e., union leaders). Approval decisions are shaped by political considerations, by the relative organizational strength of the communities submitting applications, and by

[5] We treat the per-union well budget as fixed for the purpose of constructing our primary outcome. Where we observe deviations–which we expect to be very rare–we will assess the robustness of results to dropping the affected unions.

the intensity with which local leaders press particular requests. Engineers' recommendations are therefore only partially adopted, and technically stronger sites are sometimes displaced by weaker ones. Moreover, the officials involved are multitasking agents: well allocation is one of many responsibilities, and attention devoted to improved targeting competes with other duties. Under these conditions better information alone need not translate into better allocation — it may simply be unused.

These two barriers motivate the cross-randomized design set out in Section 3, which pairs an information treatment (access to an "Arsenic Information Dashboard" presenting sub-union contamination status, allowing digitization of the year's applications, and simulating the public health consequences of alternative allocations) with an incentive treatment (a tournament awarding a discretionary "bonus" deep well to leaders whose stated allocations best prioritize high-arsenic villages).

3 Survey and Experiment Design

3.1 Overview

The experiment is a lab-in-the-field exercise embedded in the delivery of a field intervention. It is administered to Union Parishad (UP) chairpersons, henceforth "union leaders," at the moment at which the Department of Public Health Engineering (DPHE) allocation cycle for fiscal year 2026/27 is under way, and it uses the *real* applications for deep tubewell installation that the union leader's union has received in that cycle. Two barriers to arsenic-minimizing allocation are manipulated independently:

1. **Information.** Union leaders lack spatially disaggregated knowledge of arsenic contamination within their own union. We provide access to the *Arsenic Information Dashboard* (`https://nolkup.info/arrpPAP`, sample ID: `showcase`, password: `swc19`), which maps mauza-level contamination built from the Arsenic Risk Reduction Program (ARRP) testing of 6.3 million wells and the ML downscaling algorithm, allows applications to be digitized, and scores candidate allocations by the number of arsenic-exposed families they reach.

2. **Incentives.** Union leaders may not internalize the public-health return to arsenic targeting. We run a tournament in which chairmen whose selected wells go to the highest-arsenic mauzas are awarded an additional deep tubewell ("bonus tubewell") that they may place at their own discretion.

The two interventions are cross-randomized at the union level, yielding four arms (Section 3.3). In every arm the union leader is asked to select the four applications he would most like to see installed—his *preferred allocation*—from the union's real application pool. In the two arms without the tournament, incentive compatibility is instead supplied by a truth-telling device: among lottery-winning unions, one of the four wells the union leader named is installed at random, so naming a well is a costly and consequential act, but the enumerator is explicitly directed to inform the union leader that the *arsenic content* of the named wells is explicitly irrelevant to the leader's payoff.

3.2 Setting, Sample, and Respondents

Sampling frame. The frame comprises 250 unions drawn from 23 upazilas in three districts of the Chattogram division: Brahmanbaria (7 upazilas), Chandpur (3 upazilas), and Cumilla (13 upazilas). The three districts were chosen for their high levels of arsenic and for the large number of upazilas and unions that were tested by ARRP. Unions per upazila range from 5 to 16. These unions are the universe loaded into the survey instrument's cascading district–upazila–union selection lists. Of the 250 randomized unions, 200 are "active" and 50 are held as a replacement pool (Section 3.7). Active assignment is preloaded into the instrument as a server dataset (`active_200_unions_260716`) and is retrieved at the start of each interview by the union ID. Note that the ML-based downscaling method showed exceptional performance in these three district (Figure 2).

Respondent. The respondent is the union leader—one per union, and the leaders do not have discretionary authority on the allocation of wells outside of their own union. So the union is simultaneously the unit of randomization, the unit of treatment, and the unit of observation. The leader directs the upazila-level DPHE engineer on which applications to approve within his union; the upazila budget has historically been divided evenly across unions, so the leader's discretion operates over the within-union placement of a fixed number of wells. This institutional fact is what makes the union the natural unit at which to define our primary outcome.

Prevention of spillovers using credentials. The information treatment of the intervention involves providing union leaders with an access to a website that illustrates the localized arsenic contamination status to union leaders (see below for details). Although traditional conventions stop outside union leaders exercising authority on a different union leader's decisions, we further block the possibility of spillover by assigning individual IDs

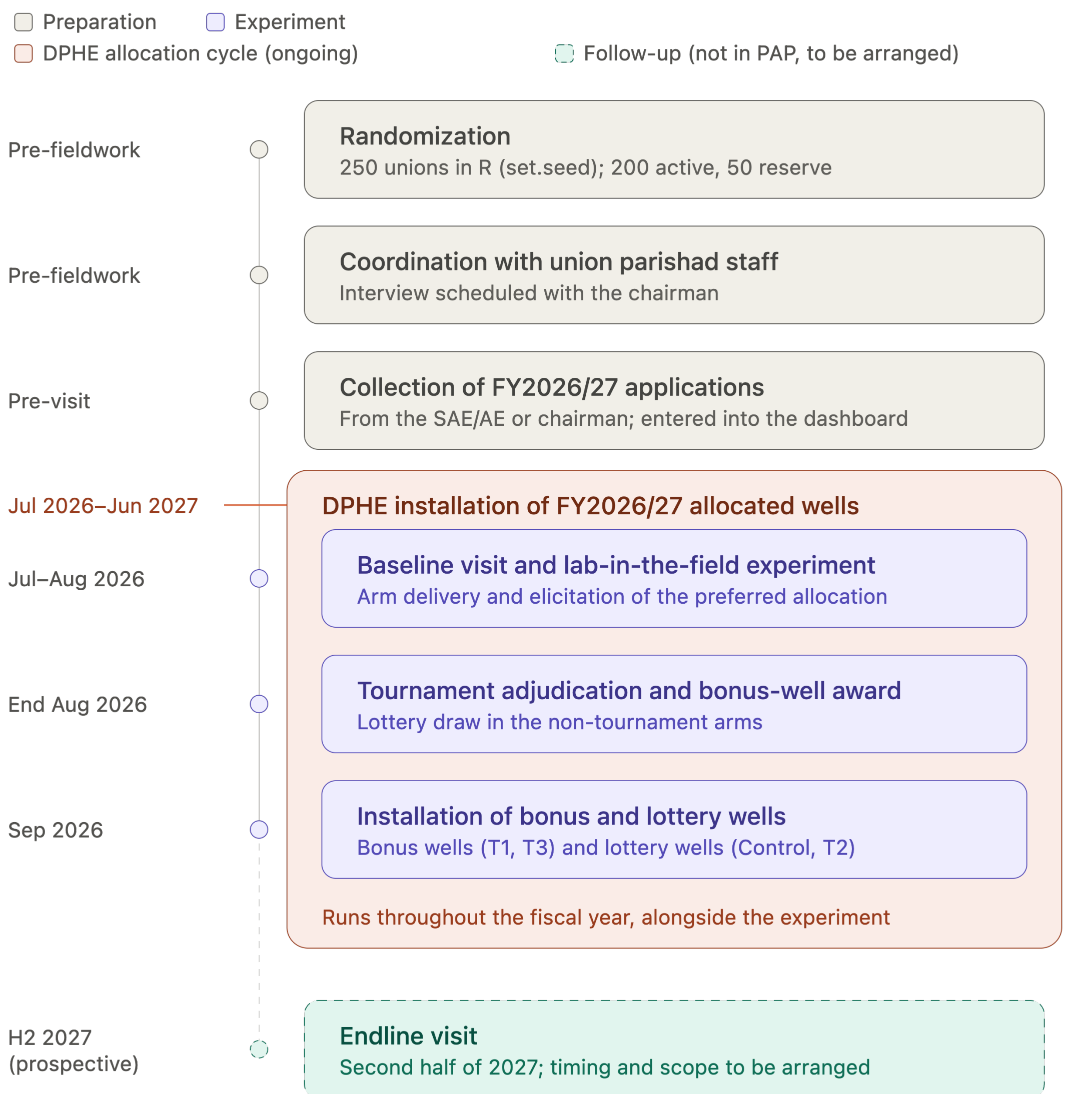


Figure 3. Timeline of intervention delivery, the lab-in-the-field experiment, and follow-up.

Notes: Randomization is conducted in *R* with a fixed seed over the 250-union frame; 200 unions are active and 50 are held as a replacement pool. Collection of FY2026/27 applications and their entry into the Arsenic Information Dashboard are hard eligibility gates. Experiment-funded wells — bonus wells in T1 and T3, lottery wells in Control and T2 — are installed following adjudication. The dashed outline denotes a prospective follow-up not pre-specified in this plan; the endline visit is anticipated in the second half of 2027, with timing and scope to be arranged. Dates in coral denote a period rather than a point: DPHE installation of the FY2026/27 allocated wells proceeds throughout the fiscal year on its own schedule and is not manipulated by the experiment. Milestones shown in gray on the spine are discrete events.

and Passwords to each union leader. The usage of each ID and password is tracked in real time and archived on a server.

Further, we note that in the lab-in-the-field environment, spillovers across union leaders are very likely immaterial. This is because there are two types of spillovers possible:

1. Treated union leaders may share their credentials with control union leaders, or influence the control union leaders' decisions. This is unlikely because we elicit the preferences *immediately*, and it is against traditional conventions for a given union leader to intervene on a different union's well allocations. The strong spatial heterogeneity in arsenic contamination across unions also means that the information from one union will be of little utility in other unions.

2. Treated union leaders may learn about the number of applications and applicants from other union leaders, upon the latter's sharing of credentials. Although this is a possibility, it is not clear how such information would affect the *preferred applications* of the union leaders.

We will, however, test for spillovers by assessing the relationship between treatment effect and the time elapsed between the first union treatment in each upazila (discussions across upazila borders are very unlikely). A positive correlation between . We also note that spillovers are a possibility for *realized*, not lab-in-the-field elicited, allocations of DPHE wells, as there are significant lags between the conferral of information and the actual decision. Although such realized outcomes are *not* part of this PAP, we will take the spillovers into consideration in the analysis of administrative data (on actual well allocations).

Eligibility. A union enters the experiment only if it satisfies two screens, all implemented as hard gates in the instrument:

1. *Real applications in hand.* The enumerator must confirm that the FY2026/27 applications were collected in advance from the upazila DPHE engineer (SAE/AE) and from the union leader himself, and that they were entered into the dashboard before the visit (`collect_realapplication`, `applicationsave_website`). If either is false the interview does not proceed.

2. *A non-trivial choice set.* The union must have strictly more than five applications, i.e. at least six (`haveenoughwells`, `howmanyapplications`). Because the union leader selects four wells, six is the minimum pool size at which the selection problem is informative. Unions failing this screen are recorded as ineligible and dropped.

We additionally record how many of the applications originate from the union leader himself (`howmanyapplications_fromUP`) and from the DPHE engineer (`howmanyapplications_fromSAE`).

These counts are pre-specified covariates: they measure the extent to which the choice set was furnished by the decision-maker.

3.3 Treatment Arms

Table 1 summarizes the 2×2 design. The information factor is the Arsenic Information Dashboard; the incentive factor is the form of the well prize (arsenic-contingent tournament versus arsenic-neutral lottery). Both factors are held fixed within a union.

Table 1. Experimental design: information × incentives

Arm	Label	Dashboard	Tournament	Prize structure
Control	`Control`	No	No	Lottery; 1 of the 4 named wells installed at random
T1	`T1`	No	Yes	Bonus well to chairmen targeting highest-arsenic mauzas
T2	`T2`	Yes	No	Lottery; 1 of the 4 named wells installed at random
T3	`T3`	Yes	Yes	Bonus well to chairmen targeting highest-arsenic mauzas

Notes: 50 unions per arm; a fifth group of 50 unions is held in reserve for replacement. In the tournament arms the bonus well is awarded on the arsenic ranking of the four *named* wells, but the union leader may site the bonus well anywhere in his union—it need not be one of the four. In the lottery arms enumerators are scripted to state explicitly that the content of the allocation does not affect the union leader's chance of winning.

3.3.1 The information treatment: Arsenic Information Dashboard

In T2 and T3 the enumerator introduces the dashboard on a tablet and delivers a scripted three-part tutorial:

1. the mauza-level arsenic contamination map (`tutorial_app_map_*`);
2. entry of the union's applications into the "simulate allocations" tab (`tutorial_app_applications_*`);
3. use of the simulator to identify allocations that maximize arsenic impact, i.e. that approach a score of 100% and the maximum number of families reached (`tutorial_app_evaluation_*`).

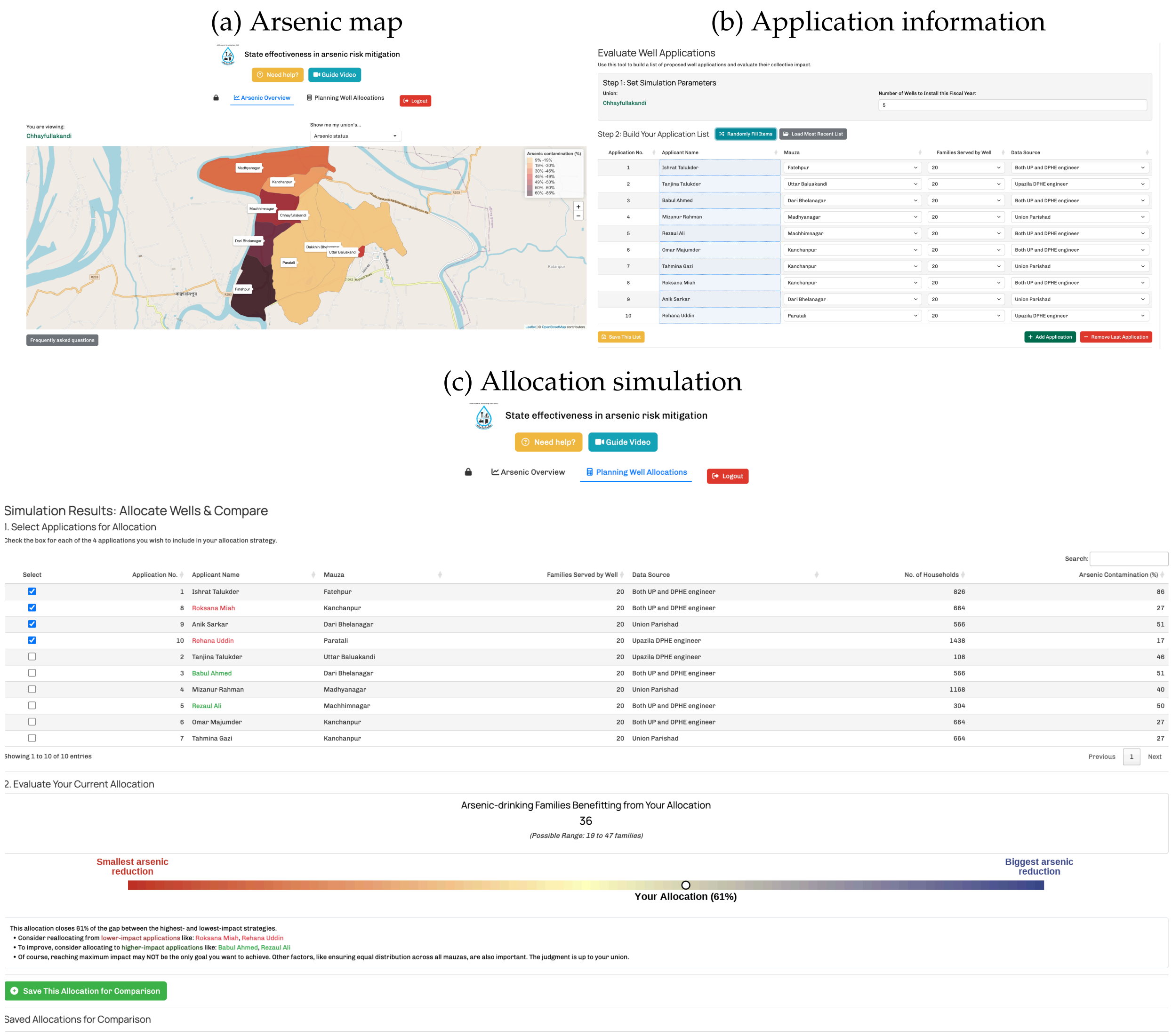


Figure 4. The Arsenic Information Dashboard shown to chairmen in T2 and T3.

Notes: Figure shows the three main features that are provided to the "information" treatment group in the T2 and T3 arms. Each union leader is given a unique ID and password code, thus not being given access to other unions' arsenic status. The example is taken for the Chhayfullakandi union, in Brahmanbaria district.

Treated chairmen also receive a printed access sheet with login credentials and a mauza-level arsenic contamination table on the reverse (`printed_infosheet`, `provideinfosheet_*`), so that the information persists beyond the visit and beyond connectivity constraints. See Appendix Figure B1 for an example. Delivery is reinforced by an official letter from the central DPHE office delivered by mail to the upazila DPHE engineer office, which establishes that the dashboard is a government-sanctioned instrument rather than a researcher artifact. The dashboard also bears the official logo of the DPHE, obatined under approval from the central DPHE office.

3.3.2 The incentive treatment: the bonus-well tournament

In T1 and T3 the introduction states that a budget of "bonus wells" exists and that the criterion for winning will be explained later in the visit; enumerators are instructed not to reveal the criterion, and not to mention arsenic, at this stage. Later—immediately before elicitation in T1, and after the pre-treatment knowledge module in T3—the enumerator discloses the criterion: whether the union leader wins depends on whether the four wells he names are sited in high-arsenic mauzas, evaluated against the DPHE-partnered mauza-level contamination dataset, with top chairmen awarded by the end of August 2026. Two features of the prize are emphasized. First, the union leader has full control over the bonus well's location. Second, the bonus well need not be one of the four wells he names. The tournament therefore rewards the *informational content* of the reported allocation without constraining the union leader's ultimate discretion, which keeps the elicitation from becoming a constrained-choice problem.

In Control and T2 the same well budget is described as "allocated wells" distributed by a lottery across participating unions, with one of the union leader's four named wells installed at random if his union wins. In T2, where the dashboard makes arsenic targeting feasible, the script is explicit and repeated: changing the allocation toward high-arsenic mauzas does *not* improve the union leader's lottery chances, and enumerators are instructed not to press the union leader to revise.

3.3.3 Credibility devices

Because both prizes are only credible if the union leader believes wells will actually be installed, three devices are used in every arm: the enumerator introduces the team as NGO Forum for Public Health working in collaboration with DPHE; the enumerator discloses that there have been letters sent from the central DPHE to the upazila's DPHE engineer office for cooperation; the union leader is given an official letter from NGO Forum restating the terms of his arm (see Appendix Section A), so that nothing material rests on recall of the verbal script (`provideletter_*`). At the close of the visit the union leader is told that the team will return at the end of FY2026/27 (`notify_willbeback`).

3.4 Blinding and Enumerator Fidelity

The arms differ in what may be said and in what order, so enumerator error is a first-order threat. Three safeguards are built into the instrument.

Deferred introduction. The form opens with an instruction not to introduce oneself yet, because the introduction is arm-specific (`note_donotintroduce`). All arm-specific content sits inside groups whose relevance conditions are keyed to the preloaded assignment.

A comprehension gate. The enumerator is shown the assignment group and the treatment-arm table, then asked independently whether this union receives the bonus-well tournament and whether it receives the dashboard and arsenic information sheet. Answers are compared against the preloaded values; only if both match (`understanding_correct`= 1) does any union leader-facing module become relevant. A mismatch triggers a correction screen that restates the union's arm and requires the enumerator to confirm having re-read the arm table before proceeding. See Appendix Figure **??** for the table that is shown to the enumerator.

Scripted disclosure order. The words "arsenic" and "website" are prohibited in the introduction of every arm. Disclosure points are fixed by the form's sequence rather than left to enumerator judgment.

3.5 Survey Flow

Table 2 gives the module order by arm. The common backbone is: enumerator checks → arm-specific introduction → union leader characteristics → digitization and display of the real application list → (arm-specific disclosures) → elicitation of the preferred allocation → stated motivations → arsenic knowledge → closing letter → interview-environment module. What varies is *what the union leader knows* when he names his four wells.

3.5.1 Union leader characteristics

Before any disclosure we collect a common battery: age, sex, highest completed education (and undergraduate field of study for degree-holders), prior post, route into office, whether the union leader was in post at the advent of the July uprising, intention to contest the next UP election, months of tenure in the post, years of prior public service, and residence and childhood mauza or union. Two features of this battery deserve note. First, the route-into-office variable distinguishes chairmen elected by their Union Parishad peers (which commonly consist of the chairman and board members–all elected from villages, except for the Secretary), elected by the general public, and assigned by the interim government via the UNO or District Commissioner—variation in the accountability channel that is specific to the current political moment and that we expect to moderate responsiveness to both treatments. Second, the residence and childhood-mauza items let us construct whether the union leader's *own* mauza is among the high-arsenic mauzas of his union, and

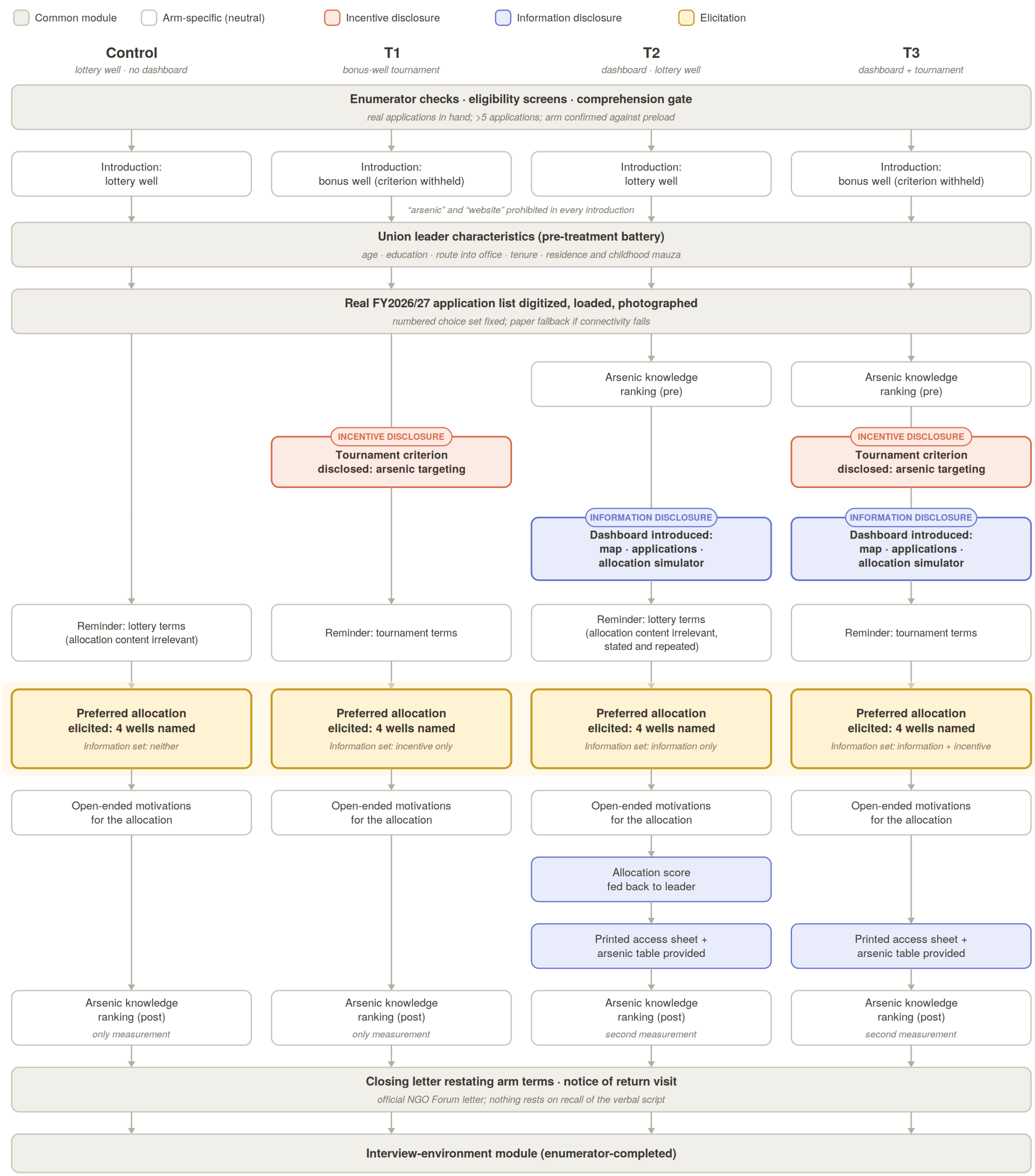


Figure 5. Flow of the lab-in-the-field experiment by treatment arm

Table 2. Module order by arm

Module	Control	T1	T2	T3
Enumerator checks; eligibility; comprehension gate	1	1	1	1
Arm-specific introduction (lottery vs. bonus well; arsenic and dashboard not mentioned)	2	2	2	2
union leader characteristics	3	3	3	3
Application list digitized, loaded, photographed	4	4	4	4
Arsenic knowledge ranking (pre)	—	—	5	5
Tournament criterion disclosed (arsenic)	—	5	—	6
Dashboard introduced; three-part tutorial	—	—	6	7
Reminder of prize terms	5	6	7	8
Preferred allocation elicited (4 wells)	**6**	**7**	**8**	**9**
Open-ended motivations for the allocation	7	8	9	10
Allocation score fed back to union leader	—	—	10	—
Access sheet / arsenic table provided	—	—	11	11
Arsenic knowledge ranking (post)	8	9	12	12
Closing letter (Appendix Section A); notice of June 2027 return	9	10	13	13
Interview-environment module (enumerator)	10	11	14	14

Notes: Numbers are within-arm order, not comparable across columns. In Control and T1 the arsenic knowledge ranking is asked once, after elicitation; "pre" therefore does not exist in those arms and the single measurement is the post measurement. In T2 and T3 it is asked twice, before and after dashboard exposure, so the pre measurement can serve as a heterogeneity dimension and the post measurement as a check that the information was absorbed. See Section 3.8 on the ordering of the T3 pre-ranking relative to the tournament disclosure.

whether his selected wells go to it, which separates targeting from self-dealing. Because this battery is administered before any treatment content, it is pre-treatment for all arms and is admissible as a covariate and as a heterogeneity dimension.

3.5.2 Construction and display of the choice set

The enumerator loads the union's real FY2026/27 applications into the dashboard and photographs the resulting numbered list; the union leader then selects by list number (Figure 4c). Where connectivity fails due to poor internet connections (`goodinternetconnection=0`), the enumerator writes the list on paper with the columns *Caretaker name / Mauza / Fam-*

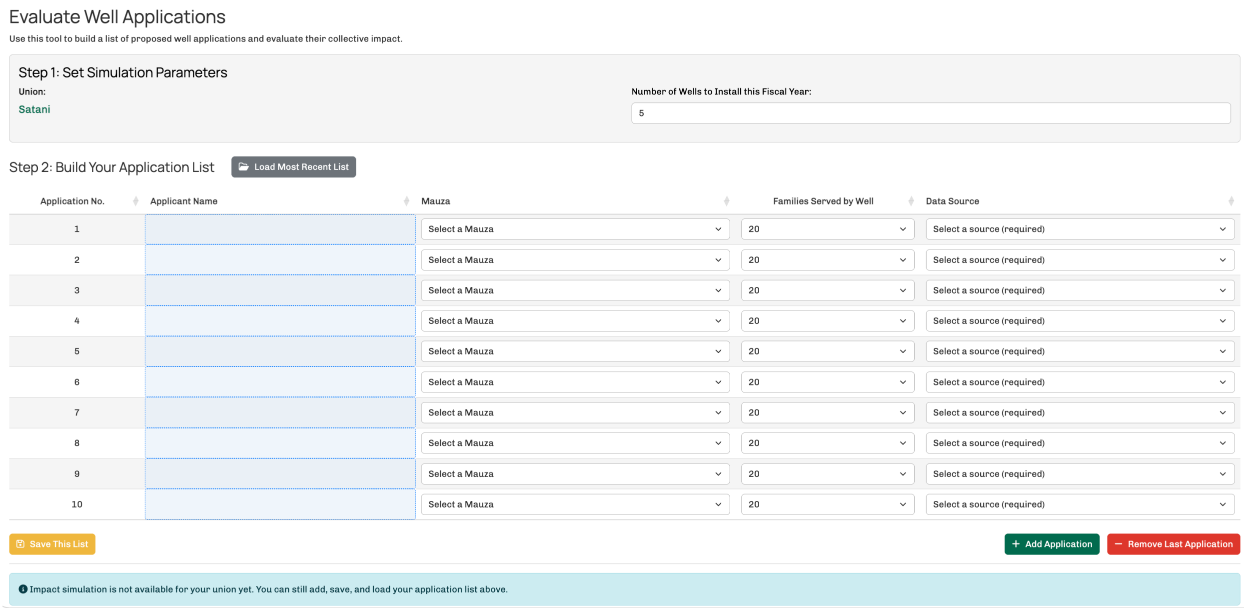


Figure 6. Dashboard website accessed by enumerators, for Control and T1 unions

ilies served by the well / Data source, photographs it, and uploads the digital record once back in coverage. The photograph is the archival record of the choice set: it fixes both the option set and the numbering used by the union leader, which is what allows us to map choices to mauzas and thus to arsenic levels ex post. Connectivity mode is retained as an implementation covariate, since in no-connectivity unions the T2/T3 union leader experiences the information as a static table rather than an interactive map.

Further, note that the dashboard is used in Control/T1 unions too, but in these unions it is only the enumerator who accesses the website. The website is only used to store and upload the collected applications (Figure 6).

3.5.3 Preferred allocation elicitation

The union leader names, in order, the first through fourth wells he would like installed (`wellchoice1--4_*`). The budget of four is set by DPHE policy for FY2026/27, so the elicitation reproduces the union leader's actual decision problem rather than a hypothetical one. Selections are recorded as application numbers from the photographed list and are merged ex post to mauza identifiers and to ARRP mauza-level contamination, from which the arsenic-exposure-reduction score of the reported allocation is computed as described in Section 4.

Immediately afterwards the union leader is asked, open-ended, what considerations entered his choices, with reasons recorded as a comma-separated list (`wellchoice_why_*`).

These will be coded into a fixed taxonomy (applicant need, water quality, village-level equity, opinions of elected representatives, opinions of unelected elders, own constituency, other)[6].

In T2 only, after the allocation is recorded, the enumerator reports the allocation's score back to the union leader, with the explicit instruction that the union leader is *not* to be asked to reach 100% (`wellchoice_infromscore`). This preserves the arsenic-neutrality of the T2 incentive while still delivering the informational content of the dashboard.

3.5.4 Arsenic knowledge

Knowledge is measured with a ranking task: the union leader is shown three mauzas from his own union—preloaded as a high, middle, and low contamination triple (`Union_Mauza_Arsenic_HighMidLow`)—and asked which is highest, which second, and which lowest in arsenic. We strictly limit the mauza pool to those that were *tested* in the ARRP campaign, and exclude those whose arsenic % levels were imputed. We can then use Kendall's τ against the true ordering as our measure of knowledge. In T2 and T3 the task is repeated with the dashboard in hand, giving a within-union leader measure of whether the information was successfully retrieved as distinct from whether it was acted upon.

3.5.5 Interview environment

The enumerator finally records the number and type of bystanders (UP secretary, other administrative subordinates, elected village representatives, other politicians, non-subordinate bureaucrats, ordinary villagers), whether the union leader engaged with all questions, interview duration, and—in T2 and T3—how interested the union leader was in the dashboard. Bystander composition is a substantive covariate, not merely a quality check: a union leader naming wells in front of the village representatives who lobbied for them faces a different problem from one answering alone.

3.6 What Each Contrast Identifies

Table 3 maps the arm contrasts to the parameters of interest. Two caveats are worth stating in advance of the results.

First, **the information contrasts bundle the dashboard with the salience of arsenic**. In

[6] There may be other taxonomies added, depending on the leaders' answers to the open-ended question.

Table 3. Arm contrasts and target parameters

Contrast	Target parameter	Held fixed
T1 − Control	Effect of arsenic-contingent incentives	No dashboard
T3 − T2	Effect of arsenic-contingent incentives	Dashboard provided
T2 − Control	Effect of information (+salience from knowledge test, see text)	Arsenic-neutral prize
T3 − T1	Effect of information (+salience from knowledge test, see text)	Tournament prize
(T3 − T2) − (T1 − Control)	Information × incentive interaction	—

T2 and T3 the knowledge-ranking task precedes elicitation, so arsenic is made salient before the union leader chooses; in Control and T1 the aresnic tests are not administred until afterwards (Table 2). T2 − Control and T3 − T1 therefore estimate the effect of the dashboard *plus* the priming induced by being asked about contamination. Because the priming is common to both information contrasts, the interaction is unaffected. Where we need to isolate the dashboard more cleanly we will lean on T3 − T1, in which the tournament disclosure has already made arsenic maximally salient in both arms and the marginal salience contribution of the ranking task is smallest. We will report the information main effect and the "dashboard plus salience" interpretation side by side rather than asserting the narrower one.

Second, **the outcome is a reported allocation, not an installed well**. The lab-in-the-field outcome measures preferences under a real but partial stake. Its external validity rests on the truth-telling and tournament devices, on the credibility devices of Section 3.3.3, and ultimately on the endline comparison with the allocations DPHE actually implements in FY2026/27, which is where the administrative primary outcome is measured.

3.7 Randomization and Stratification

Randomization is at the union level, conducted in office in R with `set.seed` for reproducibility, and assigns the 250 frame unions to five equal-sized groups: the four arms of Table 1 plus a replacement pool. The `set.seed` used is 2026.

Strata are formed from two variables: the upazila, and the maximum gains of optimization (i.e., the number of arsenic-exposed households reached under the "best allocation" minus that under the "worst allocation"). The latter is obtained from ARRP test results (well type, field-kit arsenic results, recalibrated against DPHE laboratory analysis of a random subsample). See Section 4.2 for how exactly the gains are calculated. For upazilas with at least 10 unions we split the upazila's unions at the median of the past-allocation score, giving two strata per upazila; upazilas with fewer than 10 unions form a single stratum. In the present frame 15 of the 23 upazilas have 10 or more unions, so this rule yields $15{\times}2{+}8$ = 38 strata. Stratifying on the past-allocation score matters because it is the lagged value of the outcome: unions that have historically allocated well have less headroom to improve, and unstratified assignment would leave that headroom unbalanced.

The 50 replacement unions are placed in a pre-determined random order (the `set.seed` is `20260716`). When an active union attrits—the union leader is unavailable, refuses, or the union fails the application-count screen—it is replaced by the next union in that order, and the replacement is analyzed in the arm to which it was independently assigned. We will report the number of replacements by arm and by attrition reason, and will verify balance both on the intended 200 and on the realized 200.

3.8 Implementation Notes and Anticipated Deviations

1. *Instrument version.* The design above corresponds to the SurveyCTO form *UP Chairman intervention – July 2026* (`up_chairman_intervention__july_2026`), variant "`NoTwiceQuestions`". Preloaded server datasets are `active_200_unions_260716` (arm, incentive and dashboard flags by union), `Union_Mauza_Arsenic_HighMidLow` and its long form (the three-mauza ranking triples), and `Mauza_List_For_All_Unions_CoverAllMauzas` (residence and childhood mauza lookups).

2. *Connectivity.* Where the dashboard is unreachable due to poor internet connection, T2 and T3 chairmen receive the arsenic information as the printed table on the access sheet. We will report the share of unions in this mode by arm and test whether treatment effects differ across delivery modes.

3. *Ineligible unions.* Unions with five or fewer applications never reach the chairman-facing modules. Because eligibility is determined by the union's own application flow rather than by anything we do, this screen is orthogonal to assignment in expectation; we will nevertheless report screen-out rates by arm and test for differential screening.

4. *Enumerator effects.* Enumerator identity is recorded. Because enumerators are not

randomly assigned to unions, we will report robustness to enumerator fixed effects and, given the fidelity requirements of the arm-specific scripts.

4 Outcomes

The experiment produces exactly one decision per union: the four applications the union leader names as his preferred allocation (Section 3.5.3). All outcomes below are therefore measured at the union level, with $N = 200$, and the union is simultaneously the unit of randomization, treatment and observation. This section first fixes the estimands and the estimating equation (Section 4.1), which are common to every outcome defined below; it then defines the primary outcome (Section 4.2) and the secondary outcomes (Section 4.3); and it closes with a discussion on the statistical power, calculated using data on historical allocation of wells in the 200 unions (Section 4.5). Table 4 collects the full list of outcomes.

The logic of the outcome set is as follows. The primary outcome asks whether the *allocation* improved. The secondary outcomes decompose that answer into its two proximate channels—whether the arsenic information was absorbed (knowledge), and what the union leader says he was trying to do (motivations)—and into the observable characteristics of the mauzas he chose, which let us ask whether any improvement in arsenic targeting came at the cost of other margins the literature associates with capture: wealth, population, political return, and the union leader's own village.

4.1 Empirical Specification

We will estimate the following model:

$$Y_i = \alpha + \beta_1 \mathbf{1}\{\mathrm{T1}_i\} + \beta_2 \mathbf{1}\{\mathrm{T2}_i\} + \beta_3 \mathbf{1}\{\mathrm{T3}_i\} + \delta_{s(i)} + X_i'\lambda + \varepsilon_i \tag{1}$$

where i indexes unions, $\delta_{s(i)}$ are stratum fixed effects, and X_i are the baseline characteristics of the union leader, interview environment, and pre-intervention exposure and standard errors are heteroskedasticity-robust. We will also conduct randomization inference, considering the relatively small sample size of our data. We also may pool the treatments into Information_i and Incentive_i if the interaction terms are not statistically significant.

Considering the large spread of scores that could exist in some of the revealed outcomes, we also plan to explore quantile regressions (Koenker and Bassett Jr, 1978) with the same specification (1) to study effects concentrated in the tails of the distributions of the outcomes.

Table 4. Pre-specified outcomes

Outcome	Level	Source
Primary		
Arsenic-exposure reduction score of the elicited allocation, S_u	Union	Elicitation × ARRP contamination × application list
Secondary		
Arsenic knowledge (Kendall's τ)	Union	Three-mauza ranking task
Self-described motivations for the allocation	Union	Open-ended, coded to fixed taxonomy
Mean arsenic contamination of chosen mauzas	Union	ARRP + ML downscaling
Mean household wealth, education and other observables of chosen mauzas	Union	BBS Population and Housing Census 2022
Mean population density of chosen mauzas	Union	BBS 2022
Political return to and pressure for the allocation	Union	Polling-station returns, most recent national election
Choice of the union leader's own mauza	Union	Elicitation + residence/childhood mauza items
Pre-specified but measured at endline (not analyzed in this plan)		
Arsenic-exposure reduction score of the *implemented* FY2026/27 allocation	Union	DPHE administrative records

Notes: ARRP = Arsenic Risk Reduction Program, BBS = Bangladesh Bureau of Statistics. DPHE = Department of Public Health Engineering. All outcomes are union-level; $N = 200$. The primary outcome and the mauza-characteristic secondary outcomes are constructed by merging the union leader's four named application numbers to the photographed application list (Section 3.5.2), from there to mauza identifiers, and from there to the mauza-level data sources in the third column. The endline outcome is defined here so that its construction is fixed in advance, but it is measured only after the FY2026/27 installation cycle completes and is not part of the analysis pre-specified in this plan.

4.2 Primary Outcome: Arsenic-Exposure Reduction of the Elicited Allocation

Definition. Fix a union u with application pool $\mathcal{J}_u$, $|\mathcal{J}_u| = J_u \geq 6$ by the eligibility screen, and with a budget of $k = 4$ wells set by DPHE policy for FY2026/27. Each application $j \in \mathcal{J}_u$ names a prospective well site, which we map to a mauza $m(j)$, and carries a recorded number of families the well would serve, F_j, set at 20 for default and changed upon the leader's request. The F_j is expressed as "Families served by the well" in the column of the choice set (Figure 4b). Let $\hat{\pi}_m \in [0, 1]$ denote the estimated share of drinking-water wells in mauza m that are arsenic-unsafe, constructed from ARRP field-kit readings recalibrated against DPHE laboratory analysis and, for untested mauzas, from the random-forest downscaling described in Section 2.1. The number of arsenic-exposed families reached by application j is

$$n_j = F_j \cdot \hat{\pi}_{m(j)}. \tag{2}$$

For an allocation $a \subset \mathcal{J}_u$ with $|a| = k$, the benefit is $B_u(a) = \sum_{j \in a} n_j$. Because the choice set is finite and the budget fixed, the attainable range of B_u is known exactly: writing $n_{(1)} \geq n_{(2)} \geq \cdots \geq n_{(J_u)}$ for the order statistics of $\{n_j\}_{j \in \mathcal{J}_u}$,

$$B_u^{\max} = \sum_{r=1}^{k} n_{(r)}, \qquad B_u^{\min} = \sum_{r=J_u-k+1}^{J_u} n_{(r)}. \tag{3}$$

The primary outcome is the union leader's realized allocation a_u expressed as a position in that range, in points:

$$S_u = 100 \times \frac{B_u(a_u) - B_u^{\min}}{B_u^{\max} - B_u^{\min}} \in [0, 100]. \tag{4}$$

$S_u = 100$ means the union leader selected the four applications that reach the largest number of arsenic-exposed families available to him; $S_u = 0$ means he selected the four worst. The MDEs in Table 6 are expressed in points of this score.

Why the normalization. Three considerations motivate scoring the allocation against the union's own attainable frontier rather than using the raw count $B_u(a_u)$. First, unions differ enormously in contamination and in application volume, so the raw count is dominated by variation in the choice set rather than by the quality of the choice; equation (4) conditions that variation out by construction. Second, the normalized score is exactly the quantity the union leader is being asked to move: it is bounded, it has a well-defined maximum, and the counterfactual it embeds—what *this* union leader could have achieved from *this* application pool—is the one over which he has authority. Third, the denominator of S_u is the object used to stratify randomization, since the strata are formed from the gains from optimization (Section 3.7). This makes the stratification precision-improving with repsect to the outcome, rather than merely balancing.

The normalization also has a cost worth stating in advance. In T2 and T3 the dashboard displays a score to the union leader while he chooses, and that score is built from the same formula. Our primary outcome is thus, for treated chairmen, a quantity they can see. This is intentional—the dashboard's purpose is to make the objective legible—but it means the information treatment effect on S_u combines learning about contamination with responding to a displayed target. We do not attempt to separate these two components with the present design, and we will describe the estimated information effect in those terms. The tournament effect is not subject to this concern in T1, where no score is ever shown.

Degenerate and near-degenerate denominators. Equation (4) is undefined when $B_u^{\max} = B_u^{\min}$, which occurs only if every application in the pool reaches the same number of arsenic-exposed families—in practice, when the whole union is estimated to be arsenic-free ($\hat{\pi}_m = 0$ for all mauzas in the pool). These unions do not occur in the sample. We additionally report results excluding unions in the bottom decile of $B_u^{\max} - B_u^{\min}$, where the score is a ratio of small numbers and is mechanically noisy.

4.3 Secondary Outcomes

Our secondary outcomes are the following.

1. **Arsenic knowledge.** Measured by the three-mauza ranking task of Section 3.5.4: the union leader ranks a preloaded high / middle / low contamination triple drawn from his own union, restricted to mauzas *tested* in the ARRP campaign so that the truth against which he is scored is measured rather than imputed. The outcome is Kendall's τ between his ranking and the true ordering. With three items τ takes four values, $\{-1, -1/3, 1/3, 1\}$, corresponding to zero, one, two or three concordant pairs, so we also report two coarser and more interpretable binaries: an indicator for the ranking being exactly correct, and an indicator for correctly identifying the most contaminated of the three. Measurement timing differs by arm and this is what

the outcome is for. In Control and T1 the task is administered once, after elicitation, and that single measurement is the post measurement. In T2 and T3 it is administered twice on the same triple, before the dashboard is introduced and again with the dashboard in hand. The post measurement tests whether the union leader is capable of accessing the information and updating his understanding. The pre-measurement exists only in T2 and T3 and we therefore use it as a heterogeneity dimension (see the Heterogeneity section below) rather than as a covariate in the pooled specification.

2. **Self-described motivations for the allocation.** Immediately after the four wells are named, the union leader is asked, open-ended, what considerations entered his choices. Responses are coded into the fixed taxonomy of Section 3.5.3. Each category yields an indicator, and the water-quality indicator is the one of primary interest: it tests whether the treatments change the *stated* objective and not only the chosen allocation. We report the indicators separately and a standardized index of them.

3. **Mean arsenic contamination of the chosen mauzas.** The (population-weighted) mean of $\hat{\pi}_{m(j)}$ across the four named applications. This differs from the primary outcome in dropping the families-served weighting and the within-union normalization, and it therefore isolates the contamination margin from the size margin: a union leader can raise the primary outcome either by choosing more contaminated mauzas or by choosing larger ones, and this outcome, read alongside outcome 5, separates the two.

4. **Mean socio-economic characteristics of the chosen mauzas.** Population-weighted means, across the four named mauzas, of proxies for household wealth (e.g., household structure type), educational attainment, sectoral employment (agriculture / industry / services) and religious composition, taken from the 2022 Population and Housing Census conducted by the Bangladesh Bureau of Statistics. They are the principal evidence on the distributional incidence of the treatments. Arsenic contamination and poverty are not strongly correlated, so an intervention that succeeds in raising arsenic targeting may simultaneously redirect wells away from poorer mauzas; that possibility is something we would want to detect and report, not something the primary outcome would reveal.

5. **Mean population density of the chosen mauzas.** Population-weighted mean density from the same census. Together with outcome 3, this decomposes movement in the primary outcome into a contamination component and a population component.

6. **Political return to, and political pressure for, the allocation.** Polling-station-level vote margins in favor of the incumbent Bangladesh Nationalist Party (BNP) and voter turnout in the most recent national election—the first held after the July 2024 student uprising— matched to mauzas via polling-station location, and averaged across the four chosen mauzas. Where polling-station returns cannot be obtained in practice we fall back to upazila-level results and use them as a driver of heterogeneous treatment effects rather than as an outcome, and we will state clearly which of the two is being reported. This outcome asks whether arsenic-based targeting displaces politically-based targeting, which is the mechanism the capture literature on this setting emphasizes (van Geen et al., 2016; Mobarak and van Geen, 2019).

7. **Choice of the union leader's own mauza.** An indicator for whether any of the four named wells is sited in the mauza in which the union leader resides, elicited pre-treatment in the characteristics battery (Section 3.5.1); we report the analogous indicator for his childhood mauza, the count of named wells falling in either, and the indicator restricted to the first-named well. This is the sharpest available test of self-dealing, and it interacts informatively with the primary outcome: because the union leader's own mauza may itself be high-arsenic, we additionally report the indicator separately for chairmen whose own mauza is above and below the union median contamination, which distinguishes targeting from self-dealing that happens to look like targeting.

8. **Transformations of the primary outcome**. These include:

   (a) *Unnormalized benefit.* $B_u(a_u)$, the raw number of arsenic-exposed families reached, in levels and in logs. This is the quantity a public-health planner cares about, and it is the natural scale for an aggregate statement about what the intervention would buy at national scale; it is not our primary outcome only because its across-union variance is dominated by the choice set. We may also log-transform the unnormalized benefit and the arsenic concentration levels to deal with outliers.

   (b) *Order-weighted score.* The union leader names the wells in order, first through fourth.[7] We recompute S_u weighting the named applications $4, 3, 2, 1$ and renormalizing $B_u^{\max}$ and $B_u^{\min}$ with the same weights. Under the lottery device in Control and T2, one of the four is installed at random and the order is payoff-irrelevant, so any treatment effect on the order-weighted score in excess

[7]For example, the first well is given in response to the question "which well would the chairman like to install first?" The second well is a response to the question "which well would the chairman like to install second?" and so on.

of the effect on S_u is informative about how the union leader ranks rather than only about which set he names.

(c) *Top-choice score.* The score of the first-named well alone, $100 \times (n_{j_1} - n_{(J_u)})/(n_{(1)} - n_{(J_u)})$, which isolates the single choice least contaminated by list-completion behavior.

The ordering of the secondary outcomes in the enumeration above is the descending priority order used there and is fixed by this plan.

Table 5. Shell: secondary outcomes

	(1) Knowledge (τ)	(2) Water-quality motivation	(3) Mean arsenic	(4) Wealth index	(5) Own mauza
β_1 (incentive)	—	—	—	—	—
β_2 (information)	—	—	—	—	—
β_3 (incentive and information)	—	—	—	—	—
Control mean	—	—	—	—	—
Observations	200	200	200	200	200

Notes: Table shell fixed in advance; one column per secondary-outcome family, with the remaining families (population density, political return) and the individual components of each index relegated to an appendix table of the same form. Estimation is by equation (1) with stratum fixed effects and heteroskedasticity-robust standard errors, and rows are the linear combinations in equation (**??**). Column (1) uses the post measurement, which exists in all four arms; the T2/T3 pre measurement enters only as a heterogeneity dimension.

4.4 Additional Measures Recorded, Not Pre-Specified as Outcomes

Three further sets of measures are collected and will be reported descriptively. We list them here so that their status is unambiguous: they are *not* confirmatory outcomes and no hypothesis test on them enters the multiple-comparisons hierarchy.

1. *Dashboard engagement.* Credential-level usage of the dashboard is logged in real time and archived on the server (Section 3.2), and the enumerator records how interested the union leader appeared in the dashboard. Post-visit logins are the only available evidence on whether the information persists beyond the interview. Because logins are only defined for T2 and T3, they support no across-arm test and are reported as a description of take-up.

2. *Fidelity and delivery mode.* Whether the union operated in the no-connectivity mode (Section 3.8). This enter as potential drivers of heterogeneity and robustness variables.

3. *Interview environment.* Bystander count and composition, whether the union leader engaged with all questions, and interview duration (Section 3.5.5). Bystander composition is post-treatment in the weak sense that it is recorded at the end of the visit, so we use it descriptively and, where we condition on it, we report the unconditional estimate alongside.

4.5 Statistical Power

Power is computed for the estimands of Section 4.1, measured on the primary outcome of Section 4.2 and estimated by the saturated specification of equation (1); no new specification is introduced here.

Placebo re-randomization. Rather than impose an analytic variance, we compute standard errors under the design itself. We draw the treatment assignment $K = 500$ times using the actual randomization routine over the real 250-union frame with its strata and allocation ratio, estimate equation (1) on each draw using the score implied by unions' past allocations as a placebo outcome, and average the standard error of each contrast across draws. The minimum detectable effect is then

$$\mathrm{MDE}(L) = \left(t_{1-\alpha/2,\,\mathrm{df}} + t_{\kappa,\,\mathrm{df}}\right) \cdot \overline{\mathrm{SE}}(L'b), \tag{5}$$

with $\alpha = 0.05$ (two-sided), $\kappa = 0.80$, and degrees of freedom net of the estimated coefficients and the absorbed strata. This procedure prices in the stratification, the allocation ratio and the residual variance of the actual sampling frame, which the standard $\sqrt{2(T+1)/N}$ expression doesn't easily accommodate.

Results. Table 6 suggests that the MDE for the β_1, β_2, and β_3 terms under the balanced distribution of cell sis 13.92, which is around 58.9% of the baseline SD (23.61) and 23.9% of the baseline mean (58.04).

Table 6. Minimum detectable effects by estimand and allocation

Estimand		Balanced [50/50/50/50]
β_1	(just incentive)	13.92
β_2	(just information)	13.92
β_3	(information and interaction)	13.92

Notes: Entries are minimum detectable effects in points of the primary targeting score, at 80% power and a 5% two-sided test, computed from equation (5) and averaged over $K = 500$ placebo re-randomizations of the real 250-union frame with $N = 200$ surveyed unions. Bracketed figures give the number of unions assigned to Control/T1/T2/T3.

5 Heterogeneity

We will combine two different machine learning models, the generalized random forest (GRF, Wager and Athey, 2018) and the DR-Learner (Kennedy, 2023) to flexibly estimate the conditional average treatment effect (CATE) of the treatment. We first summarize the different potential drivers of heterogeneous treatment effects, and discuss the statistical procedures used in obtaining the CATE estimates.

5.1 Potential drivers of heterogeneity

We use the following characteristics as potential drivers of heterogeneous treatment effects:

1. **Union leader's prior exposure to arsenic.** The baseline survey that accompanies the lab-in-the-field experiment collects information on multiple union leader characteristics. In particular, we hypothesize that (1) whether union leader resides in the union, (2) union leader's tenure in his position, (3) whether union leader spent his childhood in the union, (4) union leader's baseline knowledge of the mauzas' arsenic status, (5) the number of DPHE well applications possessed by the leader, and (6) the leader's education levels could be potential drivers of the leader's response to novel information on information on water quality to his decisions.

2. **Union leader's political aspirations and potential for elite capture.** As previously discussed, DPHE wells are coveted possessions that are subject to intense political pressure. Thus, we assess the heterogeneity of the information and incentives' impacts with respect to the leader's political aspirations and different aspects of the political calculus, such as (1) the union leader's plans on the upcoming union leader

election, (2) the union's (and its subsidiary mauza's) vote returns and turnouts in the most recent national parliamentary election, (3) the modality of the leader's appointment to his post (e.g., elected from a general election before the July Uprising, elected among his peers after the July Uprising), and (4) other proxies of pre-existing elite capture (e.g., the scores of prior DPHE well installations in the union, the absolute number of DPHE well installations).

3. **Survey environment.** In rural Bangladesh, union leaders work in a highly public environment where villagers and their representatives freely enter the chairperson's office with inquiries and requests. The lab-in-the-field experiment also took place within such public contexts, and we posit that the composition and number of the bystanders in the office as the survey was conducted could drive the response to changes in information and in incentives. The collected data include (1) the number of bystanders, (2) the existence of villager representatives, (3) the existence of ordinary villagers, (4) the subjective attentiveness of the union leader, and (5) the duration of the interview, in minutes.

5.2 Statistical procedures for heterogeneity analysis

5.2.1 Generalized Random Forests and R-learner

The above list presents a vector of variables X_i that could drive the heterogeneity in the treatment effect of our intervention. In practice, the different elements of X_i will likely have strong interactive and nonlinear relationship in driving the conditional expectation of the observed outcome. For example, a union leader with strong political aspirations (driver type #2) may show stronger responsiveness to the information treatment if more villagers are observing the lab-in-the-field experiment (driver type #3). Under such circumstances, machine learning based methods, such as random forests (Basu et al., 2018), can be used to flexibly model the outcome model and estimate the CATE.

Further, note that the experimental characteristic of our intervention provides us with a known propensity score. While some descriptions of the GRF/R-learner and DR-learner below (the two ML approaches that we will be using) may refer to propensity score models for completeness, we will not be estimating the propensity score model. We may, however, refer to our own computed propensity score models to understand how changes in the probability of treatment would alter the estimates of CATE and ATE in general.

The Generalized Random Forest (GRF), first developed in Athey et al. (2019), present an empirical approach for adopting machine learning methods for causal estimates. To fix ideas, we discuss the estimation approach under a binary treatment first–and then apply it to a multi-treatment arm case following Nie and Wager (2021). Our estimand of interest is CATE, defined as following under binary treatment:

$$\tau(x) = \mathbb{E}[Y_i^{(1)} - Y_i^{(0)} | X_i = x] \tag{6}$$

where $Y_i^{(1)}$ and $Y_i^{(0)}$ are the potential outcomes for union leader i, under treatment and control assignment, respectively. When $W_i \in \{0, 1\}$ represents the treatment status of union i, its CATE can be expressed as follows using the Robinson (1988) partial linearization[8] (Robins, 2004):

$$\tau(X_i) = \arg\min_{\tau} \left\{ \mathbb{E}\left([\{Y_i - m(X_i)\} - \{W_i - e(X_i)\}\, \tau(X_i)]^2 \right) \right\} \tag{7}$$

where $m(x)$ is the conditional expectation of the outcome given $X_i = x$, and $e(x)$ the propensity score (i.e., probability of treatment). In finite samples, the estimates of this $\tau(X_i)$ can be obtained from R-learner (Nie and Wager, 2021), which takes two steps:

1. Divide the data into Q folds, and estimate $\hat{m}$ and $\hat{e}$ models using machine learning-based methods. As alluded to above, we will not be training the $\hat{e}$ model in practice–as the propensity scores are already known from our experiment design. But we retain $\hat{e}$ notation for completeness.

2. Estimate the CATE by minimizing the following:

$$\hat{\tau}(X_i) = \arg\min_{\tau} \left[\hat{L}_n(\tau(X_i)) + \Lambda_n(\tau(X_i)) \right]$$

where $\Lambda_n(\tau))$ is a "penalty function" preventing overfitting, obtained over the sample $i = 1, \ldots, n$, and $\hat{L}_n$ is the mean prediction error

$$\hat{L}_n = \frac{1}{n} \sum_{i=1}^{n} \left[\left\{ Y_i - \hat{m}^{-q(i)}(X_i) \right\} - \left\{ W_i - \hat{e}^{-q(i)}(X_i) \right\} \tau(X_i) \right]^2$$

[8] To be precise, this formulation comes from the fact that the revealed outcome for a unit i with covaraite X_i and treatment status W_i can be expressed as

$$Y_i = \mu_0(X_i) + W_i \tau(X_i) + \varepsilon_i$$

Under conditional mean outcome $m(x) = \mathbb{E}[Y | X = x]$ and propensity model $e(x) = \mathbb{E}[W | X = x)$, we have the following moment condition:

$$\mathbb{E}[(W - e(x))(Y - m(x)) - (W - e(X))\tau(X) | X = x] = 0$$

where $\hat{e}^{-q(i)}$ and $\hat{m}^{-q(i)}$ are models trained on folds that the unit i does not belong to (cross-fitting).

Under multiple treatment arms (in our case with $k = 3$ potential treatments), the treatment status W can now be defined in a k dimensional space as $W \in \{0, 1\}^k$. The first step is identical, where the outcome model $\hat{m}$ and propensity score model $\hat{e}$ are trained on folds. The difference is that the outcome model $\hat{m}$ is trained on the predictor vector X_i *and* the vector of treatments $W_i \in \{0, 1\}^k$, while the propensity score model is a vector $\hat{e}(X_i) \in \mathbb{R}^k$. The second step is (Nie and Wager, 2021):

$$\hat{\tau}(X_i) = \arg\min_{\tau} \left[\frac{1}{n} \sum_{i=1}^{n} \left[\left\{ Y_i - \hat{m}^{-q(i)}(X_i) \right\} - \left\langle W_i - \hat{e}^{-q(i)}(X_i), \tau(X_i) \right\rangle \right]^2 + \Lambda_n(\tau(X_i)) \right] \quad (8)$$

where the squared bracket refers to the inner product. Note that the $\hat{\tau}(X_i)$ is a k-dimensional vector for each i, and the minimization is being conducted over the $k \times k$ dimensions.

In practice, the randomization of treatment in our context simplifies the procedure, as the propensity scores are known and $\hat{m}$ is the only quantity that needs to be estimated. We will refer to regression random forests for the $\hat{m}$ model, with tuning parameters defined using cross-validation exercises (with out-of-bag RMSE as the measure of accuracy). The random forest model will adopt honesty in node splits and leaf populations. To the extent that the GRF models do not allow categorical predictors, such predictors will be either dropped (for not ordered factor variables) or converted to numerical values, if they are ordered (e.g., education levels). The CATE will be estimated for individual union leaders, and its standard errors obtained using the out-of-bag predictions. A plot comparing the CATE estimates (with standard errors) against different X_i values and their combinations (to identify interactions) will be generated. These processes are native to the `grf` package.

5.2.2 DR-learner

The DR-learner (Kennedy, 2023) builds on the R-learner approach (Nie and Wager, 2021) to construct AIPW pseduo-outcome for individual units i. For binary treatments $W_i \in \{0, 1\}$, the AIPW pseudo outcome is

$$\hat{\varphi}(X_i, W_i, Y_i) = \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i) + \frac{W_i - \hat{e}(X_i)}{\hat{e}(X_i)(1 - \hat{e}(X_i))}(Y_i - \hat{\mu}_{W_i}(X_i)) \quad (9)$$

where $\hat{\mu}_1(X_i) = \hat{\mathbb{E}}[Y|W = 1, X_i]$ is the conditional expectation of the outcome under treatment, $\hat{\mu}_0(X_i)$ is the same but for control, $\hat{e}$ the propensity score, and $\hat{\mu}_{W_i}(X_i)$ the conditional expectation of the outcome under the observed treatment assignment. Note that the AIPW psuedo-outcome (9) is similar to the Robins partial linearization formulation (7) but has a "doubly robust" characteristic where mis-specification in *either* the outcome model *or* the propensity score model still leads to unbiased estimates.

To be concrete, the three-step process of DR-learner is (Algorithm 1, Kennedy, 2023)[9]:

1. Split the sample into two folds, (D_1^n, D_2^n), and estimate the $\hat{\mu}_1(X_i)$ and $\hat{\mu}_0(X_i)$ models for the D_1^n sample.

2. Use the D_1^n-estimated $\hat{\mu}$ models to calculate the pseudo-outcome model (9) for samples in D_2^n. Then, regress the $\hat{\varphi}$ in D_2^n sample against the covariates X. Use the regression model fitted to the D_2^n sample to obtain:

$$\hat{\tau}_{dr}(x) = \hat{\mathbb{E}}_n\left[\varphi(X_i, W_i, Y_i)\right]$$

3. Carry out the same process but using D_2^n as the data set for nuisance model estimation, and D_1^n the regression sample used for obtaining the $\hat{\tau}_{dr}(x)$ function. Take the average of the $\hat{\tau}_{dr}(x)$, one from the D_2^n model and another from D_1^n model, as the CATE estimate.

Theorem 2 of Kennedy (2023) shows that the DR-learner estimates will achieve oracle efficiency if (1) the second-stage regression estimator is "stable," and (2) either the outcome model or the propensity score model are unbiased (doubly robust).[10] The second assumption is trivially satisfied for randomized control trials.

While the mutli-treatment arm version of the DR-learner approach is not discussed explicitly in Kennedy (2023), we build on the Nie and Wager (2021) approach to the multi-dimensional treatment problem to take the following steps, where $W_i \in \{0, 1\}^3$:[11]

1. Split the sample into two folds, (D_1^n, D_2^n), and using D_1^n samples, estimate the outcome models $\hat{\mu}_k(X_i)$ for $k = \{\text{T1, T2, T3}\}$ and $\hat{\mu}_0(X_i)$ for the control samples.

2. Using the $\hat{\mu}_k(X_i)$ models from D_1^n, compute for each individual i in D_2^n sample, and for all $k \in \{\text{T1, T2, T3}\}$ the multi-treatment pseudo-outcome :

$$\hat{\varphi}_{i,k} = \hat{\mu}_k(X_i) - \hat{\mu}_0(X_i) + \frac{\mathbb{1}[W_i = k]}{\pi_k}(Y_i - \hat{\mu}_k(X_i)) - \frac{\mathbb{1}[W_i = 0]}{\pi_0}(Y_i - \hat{\mu}_0(X_i)) \tag{10}$$

[9]Here we skip discussions on the estimation of the propensity score model, as the propensity score is known.

[10]Note that in Theorem 2 of Kennedy (2023) this is expressed as $d(\hat{\varphi}, \varphi) \xrightarrow{p} 0$.

[11]For example, for the control arm the $W_i = (0, 0, 0)$, for T2 arm $W_i = (0, 1, 0)$, for T3 arm $W_i = (0, 0, 1)$.

π_k is the known propensity score (identical for all units) for treatment k, and π_0 the probability of being assigned to the control group. Run three separate regressions, one for each k, using the D_2^n pseudo-outcomes. This will yield the τ_{dr} model (using D_1^n as the source of nuisance models, and D_2^n as the source of the regression model):

$$\hat{\tau}_{dr,k}(x) = \hat{\mathbb{E}}_n \left[\varphi_k | X_i \right]$$

3. Do the same but using D_2^n as the source of nuisance models and D_1^n as the source of the regression model. This will yield another $\hat{\tau}_{dr,k}$ model; take the *average* of the two $\hat{\tau}_{dr,k}$ models as the CATE estimate for i, where the comparison is treatment k versus control (as implied in the $\hat{\mu}_0(X_i)$ term in (10)).

As in the GRF/R-learner case, we will use random forest model with real-valued predictors as the source of the mean outcome model.

References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias, "Targeting the poor: evidence from a field experiment in Indonesia," *American Economic Review*, 2012, *102* (4), 1206–1240.

Argos, Maria, Tara Kalra, Paul J Rathouz, Yu Chen, Brandon Pierce, Faruque Parvez, Tariqul Islam, Alauddin Ahmed, Muhammad Rakibuz-Zaman, Rabiul Hasan et al., "Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study," *The Lancet*, 2010, *376* (9737), 252–258.

Athey, Susan, Julie Tibshirani, and Stefan Wager, "Generalized random forests," *The Annals of Statistics*, 2019.

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Lawrence F Katz, Benjamin A Olken, and Anja Sautmann, "In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics," Technical Report, National Bureau of Economic Research 2020.

Basu, Sumanta, Karl Kumbier, James B Brown, and Bin Yu, "Iterative random forests to discover predictive and stable high-order interactions," *Proceedings of the National Academy of Sciences*, 2018, *115* (8), 1943–1948.

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On, "Predicting poverty and wealth from mobile phone metadata," *Science*, 2015, *350* (6264), 1073–1076.

Breiman, Leo, "Random forests," *Machine learning*, 2001, *45* (1), 5–32.

British Geological Survey and Department of Public Health Engineering, "Arsenic Contamination of Groundwater in Bangladesh, Volume 1: Summary," Technical Report WC/00/19, British Geological Survey, Keyworth, UK 2001.

Flanagan, Sara V, Richard B Johnston, and Yan Zheng, "Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation," *Bulletin of the World Health Organization*, 2012, *90*, 839–846.

Haushofer, Johannes, Paul Niehaus, Carlos Paramo, Edward Miguel, and Michael Walker, "Targeting impact versus deprivation," *American economic review*, 2025, *115* (6), 1936–1974.

Horneman, Alex, Alex van Geen, Dennis V Kent, PE Mathe, Yan Zheng, RK Dhar, S O'connell, MA Hoque, Zahid Aziz, M Shamsudduha et al., "Decoupling of As and Fe release to Bangladesh groundwater under reducing conditions. Part I: Evidence from sediment profiles," *Geochimica et Cosmochimica Acta*, 2004, *68* (17), 3459–3473.

Human Rights Watch, "Nepotism and Neglect: The Failing Response to Arsenic in the Drinking Water of Bangladesh's Rural Poor," Report, Human Rights Watch apr 2016.

Kennedy, Edward H, "Towards optimal doubly robust estimation of heterogeneous causal effects," *Electronic Journal of Statistics*, 2023, *17* (2), 3008–3049.

Khan, Mahfuzur, Kazi Matin Ahmed, Bidhan Dey, Mohammad Saifur Rahman, Ahasan Habib, Prabhat Barnwal, Nishat Akter Juy, Phillip P Kim, Seung Min Kim, and Alexander van Geen, "Well Water Arsenic Mitigation in Bangladesh: Benefits Outweighing Risks of Sharing Private Environmental Data," *Environmental Science & Technology*, 2025, *59* (41), 21812.

Koenker, Roger and Gilbert Bassett Jr, "Regression quantiles," *Econometrica: journal of the Econometric Society*, 1978, pp. 33–50.

Meyer, Hanna, Christoph Reudenbach, Stephan Wöllauer, and Thomas Nauss, "Importance of spatial predictor variable selection in machine learning applications–Moving from data reproduction to spatial prediction," *Ecological Modelling*, 2019, *411*, 108815.

Mobarak, Ahmed Mushfiq and Alexander van Geen, "The Politics of Poison: Elite Capture of Clean Water in Bangladesh," Working Paper S-31419-BGD-1, International Growth Centre April 2019.

Nie, Xinkun and Stefan Wager, "Quasi-oracle estimation of heterogeneous treatment effects," *Biometrika*, 2021, *108* (2), 299–319.

Olken, Benjamin A and Rohini Pande, "Corruption in developing countries," *Annu. Rev. Econ.*, 2012, *4* (1), 479–509.

Pitt, Mark M, Mark R Rosenzweig, and M Nazmul Hassan, "Identifying the costs of a public health success: Arsenic well water contamination and productivity in Bangladesh," *The Review of Economic Studies*, 2021, *88* (5), 2479–2526.

Podgorski, Joel and Michael Berg, "Global threat of arsenic in groundwater," *Science*, 2020, *368* (6493), 845–850.

Quansah, Reginald, Frederick Ato Armah, David Kofi Essumang, Isaac Luginaah, Edith Clarke, Kissinger Marfoh, Samuel Jerry Cobbina, Edward Nketiah-Amponsah, Proscovia Bazanya Namujju, Samuel Obiri et al., "Association of arsenic with adverse pregnancy outcomes/infant mortality: a systematic review and meta-analysis," *Environmental health perspectives*, 2015, *123* (5), 412.

Robins, James M, "Optimal structural nested models for optimal sequential decisions," in "Proceedings of the Second Seattle Symposium in Biostatistics: analysis of correlated data" Springer 2004, pp. 189–326.

Robinson, Peter M, "Root-N-consistent semiparametric regression," *Econometrica: journal of the Econometric Society*, 1988, pp. 931–954.

van Geen, Alexander, Kazi Matin Ahmed, EB Ahmed, Imtiaz Choudhury, MR Mozumder, BC Bostick, and BJ Mailloux, "Inequitable allocation of deep community wells for reducing arsenic exposure in Bangladesh," *Journal of Water, Sanitation and Hygiene for Development*, 2016, *6* (1), 142–150.

Wager, Stefan and Susan Athey, "Estimation and inference of heterogeneous treatment effects using random forests," *Journal of the American Statistical Association*, 2018, *113* (523), 1228–1242.

Wu, Fen, Alexander van Geen, Joseph Graziano, Kazi Matin Ahmed, Mengling Liu, Maria Argos, Faruque Parvez, Imtiaz Choudhury, Vesna N Slavkovich, Tyler Ellis et al., "Arsenic exposure reduction and chronic disease mortality," *JAMA*, 2025, *334* (23), 2104–2112.

Appendix

A Notification letters for Union Parishad Chairpersons

A.1 Control group

শ্রদ্ধেয় ইউনিয়ন পরিষদ চেয়ারম্যান,

আপনার সদয় অবগতির জন্য জানানো যাচ্ছে, “Improving State Effectiveness in Arsenic Risk Mitigation” (আর্সেনিক ঝুঁকি প্রশমনে রাষ্ট্রীয় কার্যকারিতা বৃদ্ধি) শীর্ষক গবেষণার জন্য আপনি (ইউনিয়ন পরিষদ) দৈবচয়ন পদ্ধতিতে (randomly) নির্বাচিত হয়েছেন। গবেষণা প্রকল্পটি এনজিও ফোরাম ফর পাবলিক হেলথ (NGO Forum for Public Health), জনস্বাস্থ্য প্রকৌশল অধিদপ্তর (DPHE), ঢাকা বিশ্ববিদ্যালয়, যুক্তরাষ্ট্রের কলম্বিয়া বিশ্ববিদ্যালয় এবং যুক্তরাষ্ট্রের মিশিগান স্টেট বিশ্ববিদ্যালয়ের যৌথ অংশীদারিত্বে বাস্তবায়িত হচ্ছে।

এই গবেষণার মূল উদ্দেশ্য হলো—ইউনিয়ন পরিষদ চেয়ারম্যান কিভাবে জনস্বাস্থ্য প্রকৌশল অধিদপ্তরের (DPHE) নলকূপ স্থাপনের মাধ্যমে তাঁদের এলাকার জনগণের জন্য নিরাপদ ও বিশুদ্ধ খাবার পানির ব্যবস্থা করতে পারেন, তা বিশদভাবে বোঝা। আপনি সম্প্রতি যে জরিপটিতে (survey) অংশ নিয়েছেন—সেখানে আপনার কাছে এই অর্থবছরে (২০২৬/২৭ অর্থবছর) নলকূপ স্থাপন করতে চান এমন আবেদনের তালিকা চাওয়া হয়েছে যা ইউনিয়ন পরিষদের চেয়ারম্যানদের সিদ্ধান্ত গ্রহণের প্রক্রিয়াটি বোঝার জন্য বিশেষভাবে তৈরি করা হয়েছে।

আমাদের এই গবেষণায় অংশগ্রহণ করার জন্য কৃতজ্ঞতার নিদর্শনস্বরূপ, আমরা আপনাকে একটি "বরাদ্দকৃত নলকূপ" (allocated well) প্রাপ্তিতে সহায়তা করতে চাই । আপনি যদি এই "বরাদ্দকৃত নলকূপ"টি পান, তবে আপনার পছন্দকৃত ৪ টি তালিকা থেকে দৈবচয়নের (randomly) মাধ্যমে একটি নলকূপ আমরা সম্পূর্ণ বিনামূল্যে স্থাপন করে দেব।

আমরা আন্তরিকভাবে দুঃখিত যে, সীমিত সম্পদের কারণে সকল ইউনিয়ন পরিষদকে এই "বরাদ্দকৃত নলকূপ" দেওয়া সম্ভব হচ্ছে না। তাই নিরপেক্ষতা, ন্যায্যতা ও স্বচ্ছতা নিশ্চিত করার লক্ষ্যে, কোন কোন চেয়ারম্যান এই "বরাদ্দকৃত নলকূপ" পাবেন তা নির্ধারণের জন্য একটি লটারির আয়োজন করা হবে। এই লটারিটি ২০২৬ সালের ৩১শে আগস্ট এর মধ্যে অনুষ্ঠিত হবে এবং এটি রেকর্ড করা হবে। রেকর্ডকৃত ভিডিওর সম্পূর্ণ অপরিবর্তিত (unedited) সংস্করণটি জনসাধারণের জন্য উন্মুক্ত করা হবে।

এই গবেষণা সংক্রান্ত যেকোনো প্রশ্ন বা জিজ্ঞাসার জন্য অনুগ্রহ করে এনজিও ফোরাম ফর পাবলিক হেলথ-এর গবেষণা সহযোগী জনাব মীর আবু রায়হানের সাথে (মোবাইলঃ ০১৯১২৮১০৪৭৪) যোগাযোগ করুন।

বিনীত নিবেদক,

....

মোঃ আহসান হাবিব, পিএইচডি, এনজিও ফোরাম ফর পাবলিক হেলথ

Respected Union Parishad Chairman,

For your kind information, you (the Union Parishad) have been randomly selected for the research titled ”Improving State Effectiveness in Arsenic Risk Mitigation”. The research project is being implemented in a joint partnership with the NGO Forum for Public Health, the Department of Public Health Engineering (DPHE), Dhaka University, Columbia University in the USA, and Michigan State University in the USA.

The main objective of this research is to understand in detail how the Union Parishad Chairman can arrange safe and pure drinking water for the people of their area through the installation of tubewells by the Department of Public Health Engineering (DPHE). In the survey you recently participated in, you were asked for a list of applications for installing tubewells in this financial year (2026/27 financial year), which was specially designed to understand the decision-making process of Union Parishad Chairmen.

As a token of gratitude for participating in our research, we would like to assist you in receiving an ”allocated well”. If you receive this ”allocated well”, we will install one tubewell completely free of cost, chosen randomly from your preferred list of 4.

We are sincerely sorry that due to limited resources, it is not possible to provide this ”allocated well” to all Union Parishads. Therefore, to ensure neutrality, fairness, and transparency, a lottery will be organized to determine which Chairmen will receive this ”allocated well”. This lottery will be held by August 31, 2026, and it will be recorded. The completely unedited version of the recorded video will be made open to the public.

For any questions or queries regarding this research, please contact Mr. Mir Abu Raihan, Research Associate of the NGO Forum for Public Health (Mobile: 01912810474).

Yours sincerely,

....

Md. Ahsan Habib, PhD, NGO Forum for Public Health

A.2 T1 group

সম্মানিত ইউনিয়ন পরিষদ চেয়ারম্যান মহোদয়,

আপনার সদয় অবগতির জন্য জানানো যাচ্ছে, “Improving State Effectiveness in Arsenic Risk Mitigation” (আর্সেনিক ঝুঁকি প্রশমনে রাষ্ট্রীয় কার্যকারিতা বৃদ্ধি) শীর্ষক গবেষণার জন্য আপনি (ইউনিয়ন পরিষদ) দৈবচয়ন পদ্ধতিতে (randomly) নির্বাচিত হয়েছেন। গবেষণা প্রকল্পটি এনজিও ফোরাম ফর পাবলিক হেলথ (NGO Forum for Public Health), জনস্বাস্থ্য প্রকৌশল অধিদপ্তর (DPHE), ঢাকা বিশ্ববিদ্যালয়, যুক্তরাষ্ট্রের কলম্বিয়া বিশ্ববিদ্যালয় এবং যুক্তরাষ্ট্রের মিশিগান স্টেট বিশ্ববিদ্যালয়ের যৌথ অংশীদারিত্বে বাস্তবায়িত হচ্ছে।

এই গবেষণার মূল উদ্দেশ্য হলো জনস্বাস্থ্য প্রকৌশল অধিদপ্তরের (DPHE) টিউবওয়েল স্থাপনের মাধ্যমে ইউনিয়ন পরিষদের চেয়ারম্যানরা কীভাবে নিজ এলাকার জনগণকে নিরাপদ ও বিশুদ্ধ পানীয় জল সরবরাহ করতে পারেন, তা বোঝা। আপনি যে জরিপে অংশ নিয়েছেন—সেখানে চলতি অর্থবছরে (২০২৬/২৭) স্থাপনের জন্য আপনার চাহিদানুযায়ী টিউবওয়েলের একটি তালিকা চাওয়া হয়েছে—সেটি মূলত ইউনিয়ন পরিষদের চেয়ারম্যানদের সিদ্ধান্ত গ্রহণের প্রক্রিয়াটি বোঝার জন্যই তৈরি করা হয়েছে।

আমাদের এই কার্যক্রমে আপনার অংশগ্রহণের প্রতি কৃতজ্ঞতা স্বরূপ, আমরা আপনাকে একটি "বোনাস টিউবওয়েল (Bonus well)" প্রাপ্তিতে সহায়তা করতে চাই। আপনি যদি "বোনাস টিউবওয়েলটি" পান, তবে আপনার ইউনিয়নের আওতাধীন যেকোনো গ্রামের যেকোনো সুবিধাভোগীর জন্য একটি গভীর নলকূপ (deep tubewell) স্থাপনের সুযোগ পাবেন। এই বোনাস টিউবওয়েলটি আপনার নির্বাচিত টিউবওয়েলগুলোর তালিকা থেকে হতে হবে এমন কোনো বাধ্যবাধকতা নেই, এমনকি আপনার কাছে জমাকৃত আবেদনগুলোর মধ্য থেকে নাও হতে পারে।

আপনার ইউনিয়নের সবচেয়ে বেশি আর্সেনিক দূষিত মৌজাগুলো শনাক্ত করার ক্ষেত্রে নিজেদের সক্ষমতা প্রমাণকারী শীর্ষ ৩ জন ইউনিয়ন পরিষদ চেয়ারম্যানকে এই "বোনাস টিউবওয়েল" প্রদান করা হবে। আর্সেনিকের মাত্রা সম্পর্কে আপনার ধারণা যাচাই করার জন্য আমরা ডিপিএইচই (DPHE)-এর সহযোগিতায় তৈরি করা মৌজা-ভিত্তিক আর্সেনিক দূষণ মাত্রার একটি ডেটাসেট ব্যবহার করব। জরিপ শেষে আগামী ৩১ আগস্ট, ২০২৬ এর মধ্যে আপনাদের স্কোর হিসাব করা হবে। সম্পূর্ণ স্বচ্ছতার স্বার্থে স্কোরের পূর্ণাঙ্গ তালিকা এবং ৩ জন বিজয়ীর নাম অংশগ্রহণকারী সকল ইউনিয়ন পরিষদ চেয়ারম্যানকে জানিয়ে দেওয়া হবে।

এই গবেষণা সংক্রান্ত যেকোনো প্রশ্ন বা জিজ্ঞাসার জন্য অনুগ্রহ করে এনজিও ফোরাম ফর পাবলিক হেলথ-এর গবেষণা সহযোগী জনাব মীর আবু রায়হানের সাথে (মোবাইলঃ ০১৯১২৮১০৪৭৪) যোগাযোগ করুন।

বিনীত নিবেদক,

....

মোঃ আহসান হাবিব, পিএইচডি, এনজিও ফোরাম ফর পাবলিক হেলথ

Respected Union Parishad Chairman Sir,

For your kind information, you (the Union Parishad) have been randomly selected for the research titled ”Improving State Effectiveness in Arsenic Risk Mitigation”. The research project is being implemented in a joint partnership with the NGO Forum for Public Health, the Department of Public Health Engineering (DPHE), Dhaka University, Columbia University in the USA, and Michigan State University in the USA.

The main objective of this research is to understand how Union Parishad Chairmen can supply safe and pure drinking water to the people of their respective areas through the installation of tubewells by the Department of Public Health Engineering (DPHE). The survey you participated in—where a list of tubewells as per your demand for installation in the current financial year (2026/27) was requested—was primarily designed to understand the decision-making process of Union Parishad Chairmen.

As a token of gratitude for your participation in our program, we would like to assist you in receiving a ”Bonus well”. If you receive this ”Bonus tubewell”, you will get the opportunity to install one deep tubewell for any beneficiary in any village under your Union. There is no obligation that this bonus tubewell must be from your selected list of tubewells, and it might not even be from among the applications submitted to you.

This ”Bonus tubewell” will be awarded to the top 3 Union Parishad Chairmen who prove their capability in identifying the most arsenic-contaminated mouzas in their Union. To verify your knowledge about arsenic levels, we will use a mouza-based arsenic contamination level dataset created in collaboration with DPHE. After the survey, your scores will be calculated by August 31, 2026. For the sake of complete transparency, the full list of scores and the names of the 3 winners will be communicated to all participating Union Parishad Chairmen.

For any questions or queries regarding this research, please contact Mr. Mir Abu Raihan, Research Associate of the NGO Forum for Public Health (Mobile: 01912810474).

Yours sincerely,

....

Md. Ahsan Habib, PhD, NGO Forum for Public Health

A.3 T2 group

সম্মানিত ইউনিয়ন পরিষদ চেয়ারম্যান মহোদয়,

আপনার সদয় অবগতির জন্য জানানো যাচ্ছে, "Improving State Effectiveness in Arsenic Risk Mitigation" (আর্সেনিক ঝুঁকি প্রশমনে রাষ্ট্রীয় কার্যকারিতা বৃদ্ধি) শীর্ষক গবেষণার জন্য আপনি (ইউনিয়ন পরিষদ) দৈবচয়ন পদ্ধতিতে (randomly) নির্বাচিত হয়েছেন। গবেষণা প্রকল্পটি এনজিও ফোরাম ফর পাবলিক হেলথ (NGO Forum for Public Health), জনস্বাস্থ্য প্রকৌশল অধিদপ্তর (DPHE), ঢাকা বিশ্ববিদ্যালয়, যুক্তরাষ্ট্রের কলম্বিয়া বিশ্ববিদ্যালয় এবং যুক্তরাষ্ট্রের মিশিগান স্টেট বিশ্ববিদ্যালয়ের যৌথ অংশীদারিত্বে বাস্তবায়িত হচ্ছে।

এই গবেষণার মূল উদ্দেশ্য হলো জনস্বাস্থ্য প্রকৌশল অধিদপ্তরের (DPHE) টিউবওয়েল স্থাপনের মাধ্যমে ইউনিয়ন পরিষদের চেয়ারম্যানরা কীভাবে নিজ এলাকার জনগণকে নিরাপদ ও বিশুদ্ধ পানীয় জল সরবরাহ করতে পারেন, তা বোঝা। আপনি যে জরিপে অংশ নিয়েছেন—সেখানে চলতি অর্থবছরে (২০২৬/২৭) স্থাপনের জন্য আপনার চাহিদানুযায়ী টিউবওয়েলের একটি তালিকা চাওয়া হয়েছে—সেটি মূলত ইউনিয়ন পরিষদের চেয়ারম্যানদের সিদ্ধান্ত গ্রহণের প্রক্রিয়াটি বোঝার জন্যই তৈরি করা হয়েছে। পাশাপাশি, 'আর্সেনিক ইনফরমেশন ড্যাশবোর্ড'-এর সাথে আপনাকে পরিচিত করানো হয়েছে, এটিও আপনাদের সঠিক সিদ্ধান্ত গ্রহণে সহায়তা করার জন্য তৈরি করা।

আমাদের এই গবেষণায় অংশগ্রহণ করার জন্য কৃতজ্ঞতার নিদর্শনস্বরূপ, আমরা আপনাকে একটি "বরাদ্দকৃত নলকূপ" (allocated well) প্রাপ্তিতে সহায়তা করতে চাই । লটারির মাধ্যমে আপনি যদি এই "বরাদ্দকৃত টিউবওয়েলটি" পান, তবে আপনার নির্বাচিত টিউবওয়েলগুলোর তালিকা থেকে যেকোনো একটি টিউবওয়েল [দৈবচয়নের (randomly) মাধ্যমে] আমরা সম্পূর্ণ বিনামূল্যে স্থাপন করে দেব।

আমরা আন্তরিকভাবে দুঃখিত যে, সীমিত সম্পদের কারণে আমাদের পক্ষে সকল ইউনিয়ন পরিষদের চেয়ারম্যানকে এই "বরাদ্দকৃত টিউবওয়েল" দেওয়া সম্ভব হচ্ছে না। তাই নিরপেক্ষতা, ন্যায্যতা ও স্বচ্ছতা নিশ্চিত করার লক্ষ্যে আমরা একটি লটারির আয়োজন করব, যার মাধ্যমে নির্ধারণ করা হবে কোন কোন চেয়ারম্যান এই টিউবওয়েলটি পাবেন। অনুগ্রহ করে খেয়াল রাখবেন, আর্সেনিক ওয়েবসাইট থেকে আপনার নির্বাচিত তালিকার প্রাপ্ত স্কোর কোনোভাবেই আপনার টিউবওয়েল জেতার সম্ভাবনাকে বাড়াবে বা কমাবে না।

আগামী ৩১ আগস্ট, ২০২৬ তারিখের মধ্যে এই লটারি অনুষ্ঠিত হবে এবং এর পুরো প্রক্রিয়াটি ভিডিও রেকর্ড করা হবে। স্বচ্ছতার স্বার্থে লটারির ধারণকৃত অসম্পাদিত (unedited) ভিডিওটি পরবর্তীতে সকলের দেখার জন্য উন্মুক্ত করা হবে।

এই গও ফোরাম ফর পাবলিক হেলথ-এর গবেষণা সহযোগী জনাব মীর আবু রায়হানের সাথে (মোবাইলঃ ০১৯১২৮১০৪৭৪) যোগাযোগ করুন।

বিনীত নিবেদক,

....

মোঃ আহসান হাবিব, পিএইচডি, এনজিও ফোরাম ফর পাবলিক হেলথ

Respected Union Parishad Chairman Sir,

For your kind information, you (the Union Parishad) have been randomly selected for the research titled "Improving State Effectiveness in Arsenic Risk Mitigation". The research project is being implemented in a joint partnership with the NGO Forum for Public Health, the Department of Public Health Engineering (DPHE), Dhaka University, Columbia University in the USA, and Michigan State University in the USA.

The main objective of this research is to understand how Union Parishad Chairmen can supply safe and pure drinking water to the people of their respective areas through the installation of tubewells by the Department of Public Health Engineering (DPHE). The survey you participated in—where a list of tubewells as per your demand for installation in the current financial year (2026/27) was requested—was primarily designed to understand the decision-making process of Union Parishad Chairmen. Additionally, you have been introduced to the 'Arsenic Information Dashboard', which is also designed to assist you in making correct decisions.

As a token of gratitude for participating in our research, we would like to assist you in receiving an "allocated well". If you receive this "allocated tubewell" through a lottery, we will completely freely install any one tubewell from your selected list of tubewells [randomly].

We are sincerely sorry that due to limited resources, it is not possible for us to provide this "allocated tubewell" to all Union Parishad Chairmen. Therefore, aiming to ensure neutrality, fairness, and transparency, we will organize a lottery through which it will be determined which Chairmen will get this tubewell. Please note that the score obtained for your selected list from the Arsenic website will in no way increase or decrease your chances of winning the tubewell.

This lottery will be held by August 31, 2026, and its entire process will be video recorded. For the sake of transparency, the captured unedited video of the lottery will later be made open for everyone to see.

Please contact Mr. Mir Abu Raihan (Mobile: 01912810474), Research Associate of the NGO Forum for Public Health.

Yours sincerely,

....

Md. Ahsan Habib, PhD, NGO Forum for Public Health

A.4 T3 group

সম্মানিত ইউনিয়ন পরিষদ চেয়ারম্যান মহোদয়,

আপনার সদয় অবগতির জন্য জানানো যাচ্ছে, “Improving State Effectiveness in Arsenic Risk Mitigation” (আর্সেনিক ঝুঁকি প্রশমনে রাষ্ট্রীয় কার্যকারিতা বৃদ্ধি) শীর্ষক গবেষণার জন্য আপনি (ইউনিয়ন পরিষদ) দৈবচয়ন পদ্ধতিতে (randomly) নির্বাচিত হয়েছেন। গবেষণা প্রকল্পটি এনজিও ফোরাম ফর পাবলিক হেলথ (NGO Forum for Public Health), জনস্বাস্থ্য প্রকৌশল অধিদপ্তর (DPHE), ঢাকা বিশ্ববিদ্যালয়, যুক্তরাষ্ট্রের কলম্বিয়া বিশ্ববিদ্যালয় এবং যুক্তরাষ্ট্রের মিশিগান স্টেট বিশ্ববিদ্যালয়ের যৌথ অংশীদারিত্বে বাস্তবায়িত হচ্ছে।

এই গবেষণার মূল উদ্দেশ্য হলো জনস্বাস্থ্য প্রকৌশল অধিদপ্তরের (DPHE) টিউবওয়েল স্থাপনের মাধ্যমে ইউনিয়ন পরিষদের চেয়ারম্যানরা কীভাবে নিজ এলাকার জনগণকে নিরাপদ ও বিশুদ্ধ পানীয় জল সরবরাহ করতে পারেন, তা বোঝা। আপনি যে জরিপে অংশ নিয়েছেন—সেখানে চলতি অর্থবছরে (২০২৬/২৭) স্থাপনের জন্য আপনার চাহিদানুযায়ী টিউবওয়েলের একটি তালিকা চাওয়া হয়েছে—সেটি মূলত ইউনিয়ন পরিষদের চেয়ারম্যানদের সিদ্ধান্ত গ্রহণের প্রক্রিয়াটি বোঝার জন্যই তৈরি করা হয়েছে। পাশাপাশি, ‘আর্সেনিক ইনফরমেশন ড্যাশবোর্ড’-এর সাথে আপনাকে পরিচিত করানো হয়েছে, এটিও আপনাদের সঠিক সিদ্ধান্ত গ্রহণে সহায়তা করার জন্য তৈরি করা।

আমাদের এই কার্যক্রমে আপনার অংশগ্রহণের প্রতি কৃতজ্ঞতা স্বরূপ, আমরা আপনাকে একটি "বোনাস টিউবওয়েল (Bonus well)" প্রাপ্তিতে সহায়তা করতে চাই। আপনি যদি "বোনাস টিউবওয়েলটি" পান, তবে আপনার ইউনিয়নের আওতাধীন যেকোনো গ্রামের যেকোনো সুবিধাভোগীর জন্য একটি গভীর নলকূপ (deep tubewell) স্থাপনের সুযোগ পাবেন। এই বোনাস টিউবওয়েলটি আপনার নির্বাচিত টিউবওয়েলগুলোর তালিকা থেকে হতে হবে এমন কোনো বাধ্যবাধকতা নেই, এমনকি আপনার কাছে জমাকৃত আবেদনগুলোর মধ্য থেকে নাও হতে পারে।

যে ওয়েবসাইটের (স্কোরিং অ্যাপ) সাথে আপনাকে পরিচিত করানো হয়েছে, সেখানে যাদের বরাদ্দকৃত তালিকা সর্বোচ্চ স্কোর পাবে, এমন শীর্ষ ৩ জন ইউনিয়ন পরিষদ চেয়ারম্যানকে এই "বোনাস টিউবওয়েল" প্রদান করা হবে। জরিপ শেষে আগামী ৩১ আগস্ট, ২০২৬ তারিখের মধ্যে আপনাদের স্কোর হিসাব করা হবে। সম্পূর্ণ স্বচ্ছতার স্বার্থে স্কোরের পূর্ণাঙ্গ তালিকা এবং ৩ জন বিজয়ীর নাম অংশগ্রহণকারী সকল ইউনিয়ন পরিষদের চেয়ারম্যানকে জানিয়ে দেওয়া হবে।

এই গবেষণা সংক্রান্ত যেকোনো প্রশ্ন বা জিজ্ঞাসার জন্য অনুগ্রহ করে এনজিও ফোরাম ফর পাবলিক হেলথ-এর গবেষণা সহযোগী জনাব মীর আবু রায়হানের সাথে (মোবাইলঃ ০১৯১২৮১০৪৭৪) যোগাযোগ করুন।

বিনীত নিবেদক,

....

মোঃ আহসান হাবিব, পিএইচডি, এনজিও ফোরাম ফর পাবলিক হেলথ

Respected Union Parishad Chairman Sir,

For your kind information, you (the Union Parishad) have been randomly selected for the research titled "Improving State Effectiveness in Arsenic Risk Mitigation". The research project is being implemented in a joint partnership with the NGO Forum for Public Health, the Department of Public Health Engineering (DPHE), Dhaka University, Columbia University in the USA, and Michigan State University in the USA.

The main objective of this research is to understand how Union Parishad Chairmen can supply safe and pure drinking water to the people of their respective areas through the installation of tubewells by the Department of Public Health Engineering (DPHE). The survey you participated in—where a list of tubewells as per your demand for installation in the current financial year (2026/27) was requested—was primarily designed to understand the decision-making process of Union Parishad Chairmen. Additionally, you have been introduced to the 'Arsenic Information Dashboard', which is also designed to assist you in making correct decisions.

As a token of gratitude for your participation in our program, we would like to assist you in receiving a "Bonus well". If you receive this "Bonus tubewell", you will get the opportunity to install one deep tubewell for any beneficiary in any village under your Union. There is no obligation that this bonus tubewell must be from your selected list of tubewells, and it might not even be from among the applications submitted to you.

This "Bonus tubewell" will be awarded to the top 3 Union Parishad Chairmen whose allocated list receives the highest score on the website (scoring app) you were introduced to. After the survey, your scores will be calculated by August 31, 2026. For the sake of complete transparency, the full list of scores and the names of the 3 winners will be communicated to all participating Union Parishad Chairmen.

For any questions or queries regarding this research, please contact Mr. Mir Abu Raihan, Research Associate of the NGO Forum for Public Health (Mobile: 01912810474).

Yours sincerely,

....

Md. Ahsan Habib, PhD, NGO Forum for Public Health

B Additional Figures and Tables

1. Union: Chhayfullakandi

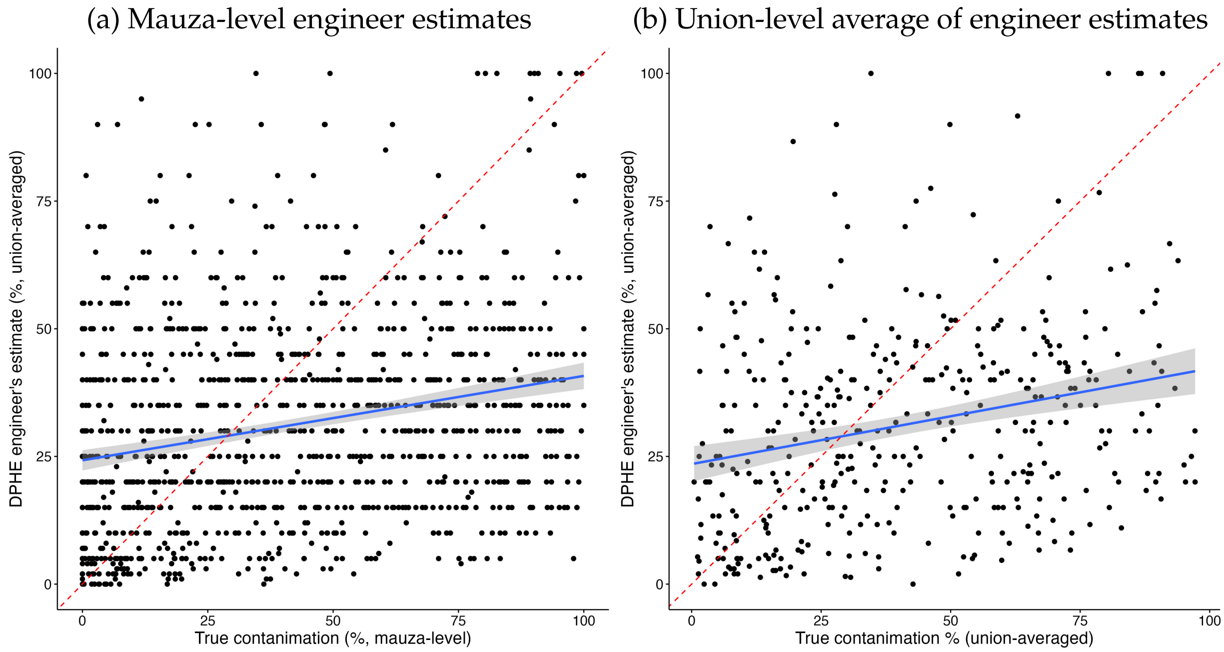
Upazila: Banchharampur

Mauza Name	Arsenic Contamination (%)	No. of families
Fatehpur	85.9	826
Dari Bhelanagar	50.8	566
Machhimnagar	50.0	304
Uttar Baluakandi	45.8	108
Madhyanagar	40.0	1,168
Kanchanpur	26.8	664
Chhayfullakandi	17.5	632
Paratali	17.1	1,438
Dakkhin Bhelanagar	9.1	1,199

Appendix Figure B1. Example of information sheet provided to a union leader

	"If you win the lottery, we're going to **build a well** in one of the mauzas that you choose. But **we are going to pick a random well among the four that you have chosen.**" ("allocation well," Truth-telling elicitation)	"If you win the lottery, we're going to give you a **bonus well** – you can do whatever you want with this. But the probability of the lottery win is increasing in how good your arsenic reduction is." ("bonus well," Incentivized elicitation)
No app	**Control [50]**	**T1 [50]**
App	**T2 [50]**	**T3 [50]**

Appendix Figure B2. Control/T1/T2/T3 arms shown to the enumerator for comprehension



Appendix Figure B3. Evaluating the DPHE engineers' understanding of local arsenic contamination

Notes: Figure assesses the accuracy of upazila-level DPHE engineer's understanding of within-upazila distribution in arsenic, who formally advise union leaders on well-water arsenic issues. Figure B3a compares the DPHE upazila-level engineers' estimates of arsenic contamination rate in different mauzas (x axis), versus their true values (y axis). The estimates are only elicited for ARRP-tested mauzas; mauzas whose arsenic contamination are gap-filled using the ML algorithm are excluded. Figure B3b averages the arsenic contamination rate (both their true values and engineers' predictions) to the union level. $y = x$ line shown in red, versus a fitted linear regression line shown in blue. A total of 146 engineers were selected for the analysis, who were elicited predictions on up to 3 mauzas for up to 5 unions in their upazila.