Pre-Analysis Plan

"Can you hear me now?": Experimental evidence on improving public service delivery through non-electoral citizen participation

Saher Asad* Asim I. Khwaja† Tiffany M. Simon‡

July 19, 2023

1 Introduction

Improving public service delivery can support economic and social development and strengthen citizen-government linkages, particularly in developing country contexts. Yet, in cases where service delivery is inadequate and citizens perceive and/or experience barriers to engagement with the government, what effective means do citizens have to hold the government representatives accountable? In developing countries, such as Pakistan, policy actors often are eager to engage with the public as can be seen by the growth in various government portals and complaint hotlines. However, the take-up of these mechanisms to reach policy actors is low and they are often too complicated for citizens to engage with at an individual level. Our project uses public education, an issue that citizens deeply care about in Pakistan, as a case in point, and we use community meetings as a means to encourage citizen interaction with policy actors to examine how these interactions can impact policy outcomes. Specifically, we use a randomized control trial that introduces community-based mobilization interventions to improve public schooling in Pakistan. Based on pilot work, we vary these interventions by: (i) policy actor type – whether citizens approach a bureaucrat directly or exert pressure through a political route and (ii) citizen gender – whether the citizens participating are women or men. In addition, for each we also include a variation in which citizens' interaction with the policy actor is more directly supported and facilitated by an NGO. We examine impact of the intervention on citizen political action, problem resolution, and school quality. Please refer to the AEA RCT Registry (AEARCTR-0011478) for full details of the intervention and study design.

*World Bank, email: sasad1@worldbank.org
†Harvard Kennedy School, email: khwaja@hks.harvard.edu
‡Princeton University Politics Department, email: tsimon@princeton.edu

2 Phase I and Phase II Study Design

A critical aspect of our study is our design of Phase I and Phase II. Phase I of our study will be conducted in 48 villages of our full 480 village sample and will allow us to refine the logistics in deploying our intervention, as well provide learning that can be incorporated into Phase II regarding the intervention, survey, and sample. We have already completed over a year of pilot activities that have informed the design of our intervention and study, thus the purpose of splitting our study into two phases is for both logistical reasons and to enable us to utilize some of the insights from the (smaller) first phase of the full study roll-out in a way that can improve the design of the final study. Specifically, by having a separate Phase I, we will have an "at-scale" pilot that will allow us to confirm our protocols and study design are working as expected in terms of the logistics of implementation and the interventions themselves as well as the survey instrument and sample size and design. Following Phase I, we will use qualitative observations of our treatment to potentially improve upon our intervention design and on implementation. We also plan to use our data to get better and more reliable estimates on our standard errors calculations which may entail then. revisiting power and sample size needs following Phase I. Note that we will not run any outcome regressions with Phase I data until we have completed the entire study. Any outcomes examined will be run with fake treatment assignment. Phase II will begin following the completion of baseline survey and intervention activities in Phase I villages. Depending on Phase I findings, we may file an amendment to our pre-analysis plan prior to Phase II should we decide to make changes to our interventions or sample.

3 Sample

3.1 Sample

Our expected full study sample will consist of 480 randomly selected villages from districts in which NRSP has worked, as further detailed below. Phase I will consist of 10% of our expected total sample of 480 villages, that is, 48 villages, randomly selected from two tehsils in the district of Khushab. This 48 village sample is stratified on school gender – in half (24) of the villages, we have selected a girls school as the government primary school of focus, while the other half (24) we will select a boys' school as the government primary school of focus. Our total sample size may change after Phase I, following analysis of standard errors.

3.1.1 Villages

Our target population is the rural poor of Punjab, Pakistan's largest state. Of the 18 districts in Punjab where our implementing partner, NRSP, works, we have selected four districts in which they have the most villages with local support organizations that also provide us with geographical spread across the region – Rawalpindi in the north, Khushab in central Punjab, and Bahawalpur and Bahawalnagar in the south.

3.1.2 Households

In each village in Phase I we will ultimately survey 8 households selected from a list-frame of 15-20 households whose children attend a selected local government primary school and are invited to participate in our intervention. While only a subset of 8 respondents will be interviewed at baseline and endline, we will invite all listed respondents to the meeting. This means we will effectively have two samples – a treatment sample, made up of individuals who are invited to participate in the intervention, and a survey sample, made up of a random subset of the treatment sample, who are surveyed at baseline. This dual sample ensures that our treatment involves an appropriate number of individuals (our pilot work suggests meetings with approximately 12 participants are most effective), while not surveying the entire treatment sample reduces our per observation cost, allowing us the budget to maximize our overall village sample. Inviting a larger group for meetings also helps intensity our effects by making sure that those aren't able to join the meetings due to day-to-day issues may be able to learn from those who are. Our power calculations on key outcomes suggest that based on ICC we only need 8 individuals per village/cluster to be able to detect sufficient effect sizes.

1. Enumerators will enter the village and locate the identified government primary school.

2. Facing the school, enumerators will identify the closest household to the school (if there is more than one household equivalently close to the school, enumerators will randomly select the first household to visit by pencil spin).

3. Enumerators will visit this household and complete a short-listing survey identifying whether the household contains a government primary school-attending child.

4. Enumerators will ask the household to point them in the direction of another household where there is a government primary school-attending child.

Using this method, we will list a total of 15 households who have a child enrolled the primary public school of interest and a parent/guardian of the gender of the corresponding village gender treatment assignment.

In addition to random direction from school, to enhance the representativeness of our sample we will also spatially drop random pins on the village map to pick at least 5 households from a random point in the village, not just those close to school or those connected to people living close to school. Using this pin drop, the enumerator will go to the pin coordinates and then approach the closest household to the pin location and then complete steps 3 and 4 again. Using this process, we intend to collect information on at least five eligible households where there is a child attending the primary school of interest and who has a parent/guardian of the gender of the corresponding village gender treatment assignment present. As part of Phase I, we will also establish if the households from the two methods are similar or different from each other in terms of meeting attendance as well as other household characteristics. This will help determine if both methods are also necessary for Phase II.

In terms of respondents, because we want to compare same-gender respondents irrespective of their treatment status, we survey both male and female respondents in 100% of the control villages and in treatment villages we will survey the respondents based on the treatment gender and will also survey both genders in 25% of male treatment villages and in 25% of female treatment villages, for a total of 6,400 respondents (2,560 respondents in control, 3,840 respondents in treatment).

A key part of our household sampling strategy is to identify the relevant study population. Our main criterion for "relevance" is being the parent or guardian of a child attending the sampled government school. However, we also want to ensure that our sample includes individuals who are likely to be exposed to our treatment (that is, individuals who attend our community meeting intervention).

To address issues of external validity, during our listing survey in Phase I, we will elicit information about a respondent's level of interest in participating in community meetings regarding their child's education in both treatment and control groups. Depending on these results in Phase I, we may decide to stratify by respondent level of interest in Phase II. This will ensure that although we may have more attendance from those who say they are interested and likely to come, we will also have attendance from some who say they were less likely to come.

Using these learnings from our Phase I, we will make improvements in our screening process and make a final call on how to best draw our sample in Phase II. That is to say, we will be able to determine how to best ensure balance between participants who are representative of parents who have children attending government primary school and are likely to take up our treatments. We will file an amendment to our pre-analysis plan prior to Phase II as appropriate.

4 Randomization

We will randomize treatment – holding of community meetings – at the village level. Within each of our four districts, we identify all villages that have both a government primary school and have ever had an NRSP local support organization. Note that NRSP is broadly active in particular districts in Punjab, so this selection ensures NRSP presence to execute the intervention. Within this set, randomization is stratified by the gender of the school in the village such that we have an equal number of girls and boys schools. The study covers four districts in Punjab, and villages in these districts will be assigned to each of the treatments as presented in Figure 1.

In total, we will assign 160 villages to pure control and 320 villages to treatment (with 40 villages in each of the eight treatment arms). Based on the power calculations, we are only interested in and powered up to detect effects between control and any 2 variations at a time (e.g. Actor x Gender, Actor x Facilitation or Gender x Facilitation).

We will conduct Phase I in the district of Khushab. Of the four tehsils (sub-district unit) in Khushab, we will randomly select two and then randomly select 48 villages that have both a government primary school and NRSP presence across these tehsils. We will then randomize treatment across these 48 villages as follows in Table 1.

Following completion and learnings from Phase I, we may make slight alterations to

Figure 1: Experimental Design

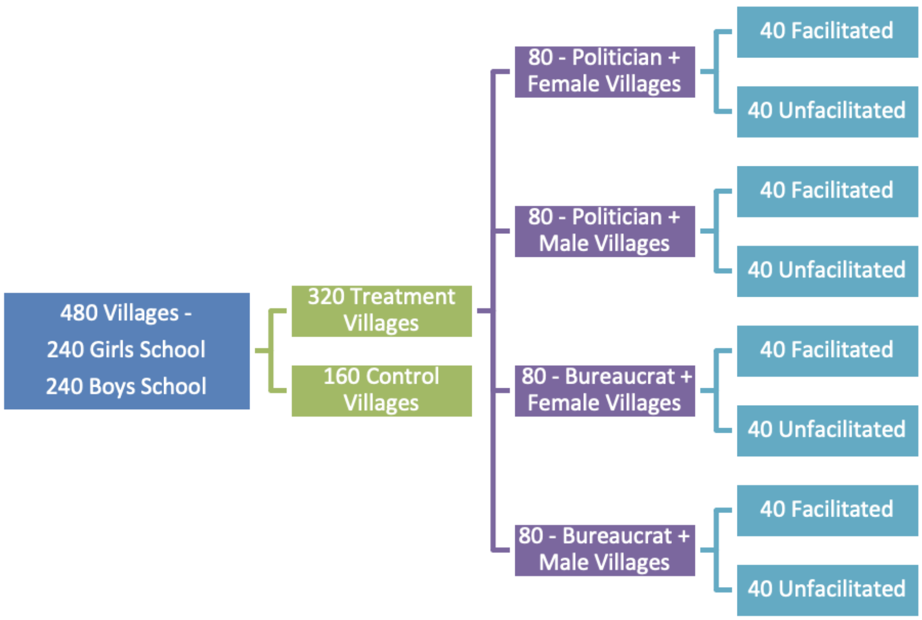

Table 1: Phase I Design

		Treatment									
		Male-Bureaucrat-Facilitated	Male-Politician-Facilitated	Male-Bureaucrat-Unfacilitated	Male-Politician-Unfacilitated	Female-Bureaucrat-Facilitated	Female-Politician-Facilitated	Female-Bureaucrat-Unfacilitated	Female-Politician-Unfacilitated	Control	Total
School gender	Boys' schools	2	2	2	2	2	2	2	2	8	24
	Girls' schools	2	2	2	2	2	2	2	2	8	24
	Total	4	4	4	4	4	4	4	4	16	48

our sampling strategy, such as change how many villages we allocate to each treatment, for example.

5 Outcomes

We will examine impact along three main types of outcomes of interest using both household level surveys and administrative data.

Research questions:

- How does citizen engagement with the government impact public service provision, namely perceptions of public service delivery?
- How does citizen engagement with the government impact perceptions of government effectiveness?
- How does this effect vary by gender and policy actor?

5.1 Primary outcomes

Our main hypothesis is that removing frictions related to coordination, information and self-efficacy will enable citizens who want to improve their children's government primary school to come together and contact policy actors and resolve the most important problems they are facing in their children's schools. In cases where we have more than one variable capturing each of the outcomes, relevant indices will be constructed. To study this hypothesis, we will investigate impacts of our intervention on the following outcomes:

1. Whether citizens take action on their education problem

2. Whether citizen action results in problem resolution

5.2 Other outcomes and channels

We also intend to explore several secondary outcomes which will highlight the different channels through which our treatments may be working via our theory of change:

1. Awareness of education problems. Citizens are often unable to act themselves because they aren't aware of the universe of problems at a school and may be unable to identify which are particularly pressing or urgent. Our intervention makes citizens aware of school performance and problems via community discussion and deliberation. Similarly, parents who are more aware of their children's education may be more compelled to take action.

2. Knowledge of and ability to contact policy actors. Even if citizens are able to identify a pressing school issue, they may not have information to act to resolve these issues in their village. Our pilot work suggests citizens lack information on who to contact and how to contact them. Our intervention thus provides citizens with information of their rights as citizens, contact information for policy actors, and guidance on how to create a message and deliver it to a policy actor.

3. Ability to take collective action. Approaching a policy actor as a group, rather than as an individual, may have greater influence on the policy actor's willingness to take action to resolve an issue. Citizens are often unable to coordinate with others in the village on education issues, and may similarly face coordination issues in approaching a policy actor in a collective manner. Our intervention helps citizens overcome the coordination issue by creating a space with a time and venue and agenda to meet and act collectively.

4. Vote choice, trust in state and self-efficacy. Following collective action, citizens who successfully reach out to a policy actor may feel empowered and self-efficacious. Positive interaction with a policy actor may increase citizen trust in state. On the other hand, if policy actors are not effective at resolving issues, citizens may be compelled to change their vote for a different party in future elections.

5. School quality. If citizens are successful in acting to solve to problems at their school, this may create a virtuous cycle that encourages them to continue to advocate for change on other school issues, increasing overall school quality.

6 Analysis

6.1 Main specification

Our basic treatment effects specification that captures the impact of our intervention is:

$$y_{vi} = \alpha + \beta T_{vi} + \theta + \epsilon_{vi} \tag{1}$$

where y_{vi} is the outcome of interest in village v at endline, measured at endline, of individual i, T_v is a treatment indicator that takes the value 1 for treatment villages, and 0 otherwise, θ theta represents randomization strata fixed effects, and ϵ_{vi} is the error term.

We will also run a specification with controls:

$$y_{vi} = \alpha + \beta T_{vi} + X_{iv}\gamma + \theta + \epsilon_{vi} \tag{2}$$

where X_{iv} is a set of control variables. We intend to separately test whether there are differences by gender, actor type and facilitation, using analogous specifications.

While our primary specification is an ITT, one can also imagine an analogous LATE specification where our treatments act as an instrument for inducing actions taken by individuals. We have two broad choices of such actions: (i) any action taken by an individual and (ii) any collective action taken by the individual. Given we are unsure how frequent either of these are in the control group and what our treatment impact will be on each, we will utilize the results of Phase I to determine which of these LATEs is relevant in our context.

We are powered up to detect combinations of any two intervention variations at any given time, including Facilitation x Gender, Gender x Actor and Actor x Facilitation. While we are not planning on running a fully interacted specification given that our sample and budget constraints imply that it is less likely that all comparisons in a fully interacted version will have sufficient statistical power, in the interest of transparency and in line with Muralidharan, Romero and Wuethrich (2022), we will also run a "long" fully interacted model:

$$y_{vi} = \alpha + \beta_1 T^U_{vi} + \beta_2 T^P_{vi} + \beta_3 T^F_{vi} + \beta_4 T^F_{vi} T^P_{vi} + \beta_5 T^U_{vi} T^P_{vi} + \beta_6 T^F_{vi} T^U_{vi} + \beta_7 T^F_{vi} T^P_{vi} T^U_{vi} + X_{iv}\gamma + \theta + \epsilon_{vi} \tag{3}$$

where T^F is a treatment indicator that takes the value 1 for villages assigned to female participation treatment and 0 for villages assigned to male participation treatment, T^P is a treatment indicator that takes the value 1 for villages assigned to politician policy actor treatment and 0 for villages assigned to bureaucrat policy actor treatment, and T^U is a treatment indicator that takes the value 1 for villages assigned to unfacilitated treatment and 0 for villages assigned to facilitated treatment.

We will also run versions of our models selecting covariates to maximize power: first, we will conduct double-lasso estimation for optimally selecting controls, identifying potential instrumental variables, and avoiding specifications search (Chernozhukov et. Al., 2016). We will thus use double-lasso to optimally select controls and improve precision by reducing standard errors.

6.2 Treatment effect heterogeneity

We will use two different methods to examine treatment effect heterogeneity. First, we intend to examine treatment effect heterogeneity on our main outcomes along the following pre-specified dimensions:

1. How compelled citizens are to take action. This is a function of interest in their children's education and how severe they find the issue.
2. How capable citizens are to take action. This is function of interest and experience in politics.
3. The gender of the school in the village.
4. Effectiveness of policy actor (based on perceptions and measures of political competition and/or bureaucratic quality)

Secondly, we will use a "split-sample approach" using a randomly selected sub-sample and machine learning (Anderson and Macgruder, 2017) to complement this analysis. In this case, we will randomly select 20% of our sample from both Phase I and II, then use machine learning to identify relevant margins of heterogeneity. Once we have identified these dimensions in our learning sample, we will examine their effect on our main outcomes in our withheld sample.

6.3 Spillover

Given the nature of our interventions, we also intend to study the potential spillovers our interventions may create. One potential way our interventions may generate spillovers is through geographical proximity. To explore this will examine outcomes in control villages in geographical proximity to the treated villages.