Identifying the Most Cost-Effective Way to Large Scale Vaccination in Rural Bangladesh

Dinesh Arora, Diwakar Mohan, Piyush Bhutani, Dai-Ling Chen, Armand Sim, Hyuncheol Bryant Kim, Debayan Prakishi, David Raitzer, Shafiun Shimul, Patrick Osewe, Asad Islam[1]

Abstract

Increasing COVID-19 vaccine uptake is of paramount importance and needs to be addressed at the earliest as vaccines are the only weapon that can end the COIV-19 pandemic. However, vaccine take-up rates are still far from universal and relatively inadequate to reach herd immunity in most countries, especially in the low-and-middle-income countries. Hence, it is crucial to understand factors that obstruct vaccine take-ups and then to identify the most cost-effective channel to address it. We conduct a cluster randomized experiment in Bangladesh involving different treatments—information campaign, information + accessibility, and information + approach by local eminent individuals—at the community level. Our sample consists of 9,090 individuals spread across 685 communities and four districts. We have completed the baseline survey and intervention. We will conduct the endline survey in September/October 2022.

JEL: I12, I14, I18, I31

Keywords: COVID-19, vaccine promotion, health behaviors, Bangladesh.

---

[1] Dinesh Arora: Asian Development Bank (ADB) (darora@adb.org); Diwakar Mohan: Johns Hopkins University (dmohan3@jhu.edu); Piyush Bhutani: ADB (piyushbhutani95@gmail.com); Dai-Ling Chen: ADB (dchen@adb.org); Armand Sim: CDES, Monash University (armand.sim@monash.edu); Hyuncheol Bryant Kim: Hong Kong University of Science and Technology (hbkim@ust.hk); Debayan Prakishi: IITK and Indian Statistical Institute Kolkata (pakrashide@gmail.com); David Raitzer: ADB (draitzer@adb.org); Shafiun Shimul: University of Dhaka (shafiun.ihe@du.ac.bd); Patrick Osewe: ADB (posewe@adb.org); Asad Islam: Monash University (asadul.islam@monash.edu).

1. Introduction

More than six million deaths have been attributed to COVID-19 pandemic (as of June 2022).[2] The reported case numbers are likely to be lower in many developing countries than in the developed countries, which is more of a reflection on the lower state capacity in the former than in the latter. The actual case numbers and adverse consequences of the pandemic are arguably higher in the developing world than that in the developed countries. Despite the gap in case numbers, the differences in relative death counts have arguably become smaller after the vaccines are in place and administered at large scale. A recent study estimates that vaccines prevented global 20 million excess deaths due to the COVID-19 pandemic (Watson, et al., 2022).

Nevertheless, even with a decent vaccination coverage, the world has not been able to fully control COVID-19 transmission across the globe. After having been consistently declining since March 2022, the global cases—attributed to the emergence of the new coronavirus variant—have started rising again in June 2022. Developing countries, including Bangladesh, are not exceptions. Vaccination rates in Bangladesh have been arguably high—the first and second dose vaccination rates have reached 78 % and 72% (as of July 2022), respectively—but they have plateaued since April 2022. Moreover, the daily cases have now averaged roughly 1,000 between June and July 2022, a significant jump from the daily cases in the previous months (March to May 2022) which are less than 100 in most cases. Together, these suggest that large-scale vaccination is still paramount in effort to reduce harmful consequences of COVID-19. Motivating people to get vaccines and, as an extension, practice health-enabling behaviors is more important than ever.

[2] https://covid19.who.int/

We address these issues in this project. One first needs to identify what are the barriers to vaccination and which policies work and do not work. These pose a great challenge because policy makers in low-and-middle-income countries face complex and multifaceted problems in scaling up vaccination: prevalent vaccine hesitancy and uneven access to vaccines supply in rural and remote areas (Reza et al., 2022). In this project, we test the effectiveness of interventions to address hesitancy and access related issues through a randomized controlled trial. The main goal is to increase confidence in the minds of policymakers who are entrusted with scale up and implementation.

We conduct a cluster randomized experiment that implements various treatments in different locations (rural and urban) in four districts—Khulna, Satkhira, Chandpur, and Laxmipur—in Bangladesh. We provide several different treatments that attempt to address two important issues that might impede vaccination progress in Bangladesh. First, misinformation and misconception about vaccines. Bangladesh is one of developing countries that shows high prevalence of vaccine hesitancy, around 30-40%, as suggested by recent studies. One study estimates that, out of 1,134 individuals, about one-third of them, mostly males, over the age of 60, unemployed and from low-income households were reluctant to have vaccines (Ali and Hossain, 2021). Another study estimates the prevalence of vaccine hesitancy at 41% with women, Muslims and those living in the city corporation areas most hesitant (Hossain et al., 2021). To address these concerns, we provide evidence-based information related to COVID-19 vaccines—collected from local sources, e.g., ministry of health, and global sources, e.g., WHO—to all treatment groups. Another group receives an additional treatment: encouragement and information campaign delivered personally in their homes by local eminent individuals in their villages.

Second, issues on access to vaccines. Since at the beginning, it was mandatory to complete online registration for vaccination, and vaccination was given only in limited facilities, people with less access to technology and geographical limitation, and low literacy worked as barriers to access to vaccination. To address these issues, randomly selected participants receive help with vaccination enrolment, transportation to a vaccination center, and reminders to vaccine take up.

In total, we have three treatment groups—*information campaign only, information + accessibility,* and *information + local ambassadors*—and one control group—that does not receive any treatment. Our final sample comprises 9,090 individuals spread across 685 communities in four districts. We will conduct the endline survey in September 2022.

This study is related to the growing literature on the effects of nudging and information provision on health behaviors and COVID-19 vaccination decisions. Recent studies show that information provision through various media successfully affect health behaviors during the pandemics: social distancing, handwashing, and masking. Banerjee et al. (2020) find that households in West Bengal that received links to video on information about COVID-19 show significantly more adherence to social distancing, handwashing, and hygiene behaviors. Siddique et al. (2020) also find significant impacts of sending SMS plus phone calls—as opposed to receiving SMS alone—to respondents in Uttar Pradesh and Bangladesh. Another means of effective information campaign intervention to increase vaccination is conducted through an influential local leader working as a vaccine "ambassador" (Banerjee et al., 2019). In this study, we complement the growing literature in this area using local influential individuals in a village—identified by local village leaders—to serve as a COVID-19 vaccine ambassador following Banerjee et al. (2019).

This study complements the literature on the effects of improving access and knowledge through personal home visits on health behaviors and vaccine take-up. Kim et al. (2017) show that home visit programs promote HIV-testing rate by 55 percentage points. Door-to-door campaigns are also effective in influencing tetanus vaccine hesitant child-bearing women in rural Nigeria (Sato and Takasaki, 2021). Improving access to a vaccination site increases chances of an individual getting COVID-19 vaccines by 4 to 12 percentage points (Khan et al., 2021). We complement this literature by providing assistance to registration for vaccination and transportation.

Overall, this study, to our knowledge, is among the first large-scale field experiments that test several strategies to promote COVID-19 vaccines that involve information, door-to-door information campaigns by local eminent individuals, and improvement in access to vaccines in a developing country context. A large proportion of studies in this literature are mostly concentrated in developed countries, such as in Germany (Klüver et al., 2021) and the United States (Chang et al., 2021), where their citizens are relatively richer and more educated than those in developing countries. It has not been examined whether responses to interventions are generalizable to developing countries. Most importantly, all of the interventions included in study are expected to be low cost, and hence, if any of them are effective, then some or all can be implemented at scale. Therefore, documenting evidence of these effects in developing countries is important because it can provide inputs for policy makers in devising effective strategies to increase vaccination rates.

2. Research Design

We conduct a cluster randomized controlled trial where the interventions are randomly allocated at the community (village) level. The setting of this study takes place in villages in four districts: Khulna, Satkhira, Chandpur, and Laxmipur.[3] We chose four districts: Khulna, Satkhira, Chandpur, and Laxmipur for two reasons: 1) its relatively

[3] The list of these villages was obtained from the local administration offices in each district. As such, we were unable to randomly select all participants for this study.

lower uptake when we initiated the study and 2) its relative distance to Dhaka—Chandpur and Laxmipur are relatively close from Dhaka whereas Khulna and Satkhira are relatively far from Dhaka.

2.1. Randomization Procedure and Sample Selection

The randomization process is conducted in two stages. First, we randomly selected 685 villages —out of our pool of villages—to be allocated into treatment and control groups. In addition to obtaining administrative data on number of villages in each district, we also use our network of field workers and informants in the selected villages to collect baseline information on village characteristics, such as proportion of Muslims, distance to the nearest COVID-19 centers (in km), and distance to the nearest various levels of schools (in km). For convenience purpose and prevention of treatment effects spillover, we imposed several restrictions on the selected villages: (i) villages are not located too far from district capitals (in km) and (ii) villages are located at least 2 km apart from each other.

Next, we proceed to select individuals. To do so, we first obtained a list of households whose members remained unvaccinated from village heads and/or officials. This list officially does not exist at the village level. In absence of the list, enumerators had to rely on the snowball approach, in which participants recommended the next potential participants.[4] When the list contains more than 13 individuals, we randomly selected 13 individuals from that list.[5] To obtain valid survey responses, we imposed some exclusion restrictions for potential participants. For example, we exclude an

[4] Enumerators also used the snowball approach when the list contains less than 13 individuals per village, which is the target number for this study.

[5] In some villages, depending on the availability of eligible participants, we selected more than 13 participants while in other villages we only managed to select less than 13 participants. But, on average, we managed to recruit 13 participants per village.

individual from our sampling framework if he is in a condition that does not allow him to understand and respond to interview questions—individuals who are extremely sick.

The final study sample comprises 9,090 individuals from 685 communities (rural and urban) spread across four districts or about 13 individuals per community on average. We randomized these communities into three treatments and one pure control group. In total, we have 137 communities (1,791 individuals) randomly assigned to *Pure Control,* 137 communities (1,948 individuals) randomly assigned to *Information Campaign Only,* 206 communities (2,628 individuals) randomly assigned to *Information + Ambassador*, and 205 communities (2,723 individuals) randomly assigned to *Information + Accessibility.* See Figure 1 for treatment groups and allocation of participants.

2.2. Intervention

Information campaign only. Participants in this group receive a set of information about misperceptions on COVID-19, the available vaccines, distribution of infection and fatality rates, among others. We collected the information from the guidelines provided by the WHO, national health ministries, and local governments. The 1-page information sheet is conveyed verbally by the field workers for about 10-15 minutes.

Information campaign and accessibility. Participants in this group receive information and free assistance related to accessing vaccines. They receive help with registration/enrollment, information on the nearest vaccination centers, travel services/costs, and brief reminders/encouragement via brief phone calls and text messages (SMS) roughly two weeks after completion of the intervention period. Staff of our NGO partner, GDRI (Global Development Research Institute), are assigned to deliver the intervention.

Information campaign and campaign/motivation by local eminent figures. Participants in this group receive information and encouragement delivered by local eminent figures (we refer to them as vaccine ambassadors) via personal home visits. To recruit the ambassadors, we approached village leaders (village heads or elders) to identify the most respected and trusted eminent figure in their community (e.g., religious leaders, teachers, doctors) and ask him/her to propagate information to address misconceptions related to the available COVID-19 vaccines.

Specifically, the selected ambassadors are asked to disseminate specific information (in addition to basic information being provided to all) to address misconceptions related to the available vaccines and convince them to get vaccinated. The ambassadors will pay a personal visit to participants' houses once during the campaign period. The meeting will last for 60 minutes at most. To minimize the risk of infection, the ambassadors follow strict health protocols (e.g., every respondent and ambassador will wear a mask, maintain 1.5 meters between individuals, open air space), and try to convince them to get vaccinated.

To establish credibility, the selected ambassadors should have received at least one dose of vaccines. Prior to the intervention, the ambassadors attended offline training and information sessions about COVID-19 vaccines.

Control group. Participants in this group—the 'pure' control group—do not receive any treatment. For analysis, we also include the *Information Campaign Only* group as an additional control group. Thus, the control group consists of *Pure Control* and *Information Campaign Only*.

Timeline. The study is expected to run between November 2021 - December 2022. We conducted the baseline survey from February 2022 until May 2022. Using baseline data, we then randomized the locations and respondents into control and treatment groups. Intervention started in June 2022 and ended in August 2022. The research team will give reminders to the intervention group participants and verify their vaccination and registration status after the intervention period and before the endline survey. They will be revisited one month after the verification for the endline survey, scheduled between September and October 2022.[6]

3. Data Collection

The main data sources in this study come from the baseline survey—completed in mid-June 2022—and the endline survey—will be collected in September-October 2022.[7] We will also use administrative data to verify vaccination status of participants.

3.1. Baseline survey

In the baseline survey, we collected variables that may predict vaccination decisions, such as basic socio-economic characteristics, compliance to health protocols related to COVID-19, morbidity history, and beliefs as well as attitudes toward COVID-19 vaccination. Below we highlight a set of information we collected during the baseline survey.

1. Basic socio-economic characteristics, such as gender, age, educational attainment, employment status, marital status, and total expenditure
2. Concern about COVID-19 for themselves, their family and their friends: risk of infection and sickness, risk of being a confirmed close contact, risk of quarantine
3. Attitude and accessibility toward COVID-19 vaccine

[6] Conditional on additional funding, we plan to have an additional survey in mid-2023 to verify booster vaccine status.

[7] We will also conduct midline survey soon after the intervention in July only to collect information on vaccination take up and registration status. Endline survey is conducted one month after the midline survey.

a. Attitude or opinion towards private benefits of vaccination, such as lower risk of falling sick, lower risk of infecting others, freedom of movement

b. Accessibility: barriers to access vaccines.

4. Compliance to health protocols related to COVID-19

a. Handwashing, mask wearing, maintaining physical distance of 1,5 meters from other people outside the house, avoiding close physical contact, or shaking hands with other people who do not live in the same household, among others.

5. Beliefs and knowledge about the COVID-19 vaccines

a. Knowledge about lower risk during future COVID-19 waves, lower risk of becoming a confirmed close contact, economic recovery/employment growth, ability to socialize, ability to gather in public, etc.

6. Vaccination status of other family members within the household and in the neighborhood.

3.2. Hypotheses

We hypothesize that the intervention:

1. Increases vaccine uptake and intention of participants and their relatives/families/friends (spillover treatment effects).

2. Improves knowledge and beliefs about COVID-19 vaccines.

3. Reduces the median time to vaccination (between the intervention and the endline survey periods).

4. Changes compliance to health protocols related to COVID-19. The direction could be positive or negative. It is hard to predict because it could go in either direction depending on other factors, such as individual’s risk preferences and policy changes.

5. Improves mental health status.

More importantly, we will also investigate which treatment is the most effective in increasing vaccine uptake and intention, but we do not have hypothesis on this matter.

3.3. Outcome Variables

3.3.1. Primary Outcomes

The main goal of this study is to investigate which treatment is the most effective in affecting vaccination decision, measured by vaccine uptake and intention.

Vaccine uptake. This outcome is an indicator for whether participants take up at least one dose of COVID-19 vaccine. To confirm vaccination status, the enumerators will verify physical or digital vaccination proof recognized by the government. Additionally, we are also interested in examining whether the intervention has any impacts on how fast participants take up vaccines after receiving the intervention. This will be measured by response time to vaccination.

Vaccination intention. In addition to vaccine uptake, we will also evaluate vaccination intention as an alternative primary outcome for those who have not taken up vaccine. *Vaccination intention* measures how willing an individual is to get vaccinated, which is constructed from the following question: "*Currently, how willing are you to be vaccinated against COVID-19?*" Participants give responses in 1-5 Likert scale range, where 1 refers to extremely unlikely and 5 refers to very surely. We will normalize these responses to have support between 0 and 1. In addition, we also use an indicator variable for whether participants report that they already registered and/or visited a COVID-19 vaccine center but could not get vaccinated due to health, e.g., high blood pressure, or other reasons.

Vaccination intention is a commonly used outcome to measure vaccination decision in the literature (e.g., Campos-Mercade et al.,2021; Chang et al., 2021; Klüver et al., 2021). Measuring the effects on vaccination intention is crucial as there might be some participants who are declined by health workers at the vaccination centers due to some health reasons (e.g., high blood pressure). When this happens the treatment effects are likely underestimated.

Vaccination status of others. In addition to participants' vaccination status, we will also investigate possible spillover effects of our intervention by examining the effects of the intervention on self-reported vaccination status of other household members, and neighbors.

Because we measure several closely related outcomes of vaccination decision, we will correct for multiple hypothesis testing problem following Anderson (2008). We will apply the same method to outcomes related to *vaccination status of others.*

3.3.2. Secondary Outcomes

We will measure other outcomes that may be affected by the intervention: health behaviors and mental health. Health behaviors are measured by compliance to COVID-19 health protocols, such as hand washing and mask wearing, and COVID-19 infection, an indicator for getting infected with COVID after intervention period. We use two variables to measure the state of mental health: general mental health and mental health attributed to COVID-19.

Compliance to COVID-19 protocols. This outcome is an index variable derived from survey responses regarding compliance with COVID-19 health protocols such as handwashing, mask use, and physical distance. Compliance is assessed using two types of responses: binary (yes/no) and Likert scale (1-4). These responses are used to generate two indices. To construct the index for binary responses, we assign 1 if one responds 'yes' to each activity and take the average value of all responses. To create the index, we first normalize each Likert-scale response to have support ranging from 0 to 1 and then take the average value of all responses.

COVID-19 infection. This outcome is an indicator variable that equals to one if respondents report getting infected with COVID after the intervention.

Mental health. We use two measures of mental health: general mental health and mental health attributed to COVID-19. Responses to general mental health question are measured using Likert scale where 1 refers to rarely or not at all (≤ 1 day) and 4 refers to often (5-7 days). Responses to mental health attributed to COVID-19 are measured using Likert scale where 1 refers to strongly disagree and 5 refers to strongly agree. These responses will be normalized to have support between 0 and 1 and we will take the average value of all responses to create the index.

3.3.3. Intermediate Outcomes

We will explore the effects of the intervention on some intermediate outcomes to investigate possible channels.

Knowledge and beliefs about COVID-19 and vaccines. This variable reports a participant's responses to questions on knowledge about COVID-19 and its prevention, severity of COVID-19 infection, benefits of COVID-19 vaccines, among others. Responses are measured using Likert scale where 1 refers to strongly disagree and 5 refers to strongly agree. We will construct index variable from these responses for each domain (e.g., knowledge about COVID-19). To do so, we will first re-code all variables so that higher values correspond to the same direction (e.g., questions on severity of COVID-19), next we normalize these responses to have support 0 to 1 and take the average of all responses to construct the index of each variable. Because the whole index comprises of several domain indices, we will correct for multiple hypothesis testing issue following Anderson (2008).

Quality of the information intervention (subjective assessment). This variable records participants' subjective assessment over the quality of the information delivered by field workers and participants' assessment on how convincing the field workers were in disseminating the information. We will normalize the Likert-scale responses—1 refers to least convincing and 5 refers to most convincing (for questions 1 and 2) and 1 refers to very bad and 5 refers to very good—to these questions to have support between 0 and 1. Because this outcome comprises of several closely related variables, we will create an index that corrects for multiple hypothesis testing problem following Anderson (2008).

3.4. Summary Statistics and Balance Test

Table 1 reports the summary statistics of baseline individual and village characteristics as well as balance tests between treatment groups. Panels A and B show individual and village characteristics, respectively. Column 1 shows total number of observations. Column 2 reports the average value of baseline characteristics of all respondents. The average respondent is relatively young, about 27 years old, and majority of respondents are female, 70 %. Even though many of them are unemployed (78 %) and completed lower-than-secondary level education (69 %), a large proportion own a house (97 %) and only 12 % receive government assistance.

Columns 3 to 8 report coefficients from regressions of each baseline on treatment group indicators. We show that only 4 out of 156 coefficients across balance tests are statistically significant at the 5% or 10 % level. We also perform joint orthogonality tests to evaluate an overall balance between groups across all baseline variables. Overall, these tests demonstrate that our randomization is successful in creating balance across treatment groups.

4. Empirical Analysis

4.1. Main Results

We estimate the following regression specification

$$Y_{i,t=1} = \alpha + \beta_1 INFO_i + \beta_2 INFOACCESS_i + \beta_3 INFOLEADERS_i + \delta Y_{i,t=0} + X'_{vi}\tau + \varepsilon_i \quad (1)$$

where $Y_{i,t=1}$ denotes vaccination decisions (take-up and intention) and other outcomes of individual *i* in the endline survey. $INFO_i$ is an indicator for respondents that are assigned to the *information campaign only* group. $INFOACCESS_i$ is an indicator for respondents that are assigned to the *information + accessibility assistance* group. $INFOLEADERS_i$ is an indicator for respondents that are assigned to the *information + influential individuals (leaders) ambassadors* group.[8] We include a vector X_{vi} of individual covariates (age, index of compliance to COVID-19 protocols, and indicators for being male, being married, being Muslim, living in rural areas, living in joint family, owning house, government assistance beneficiary, completed secondary-level education, being employed, and living in a high monthly income household) and village covariates (proportion of Muslims, nearest distance to COVID-19 centers (in km), nearest distance to community clinics (in km), nearest distance to railway stations (in km), nearest distance to secondary schools (in km), nearest distance to colleges (in km), nearest distance to post offices (in km), nearest distance to banks (in km), nearest distance to police stations (in

[8] Pure control group serves as the baseline group.

km), and nearest distance to hospital/doctor (in km)). We will use double LASSO (Belloni et al., 2014) to objectively select baseline covariates as a robustness test.

We will also include the baseline value of outcomes whenever available (e.g., vaccination intention). Standard errors ε_i are clustered at the village level.

Our main coefficients of interest $\beta_1, \beta_2, \beta_3$, capture the intention to treat (ITT) effects. We will compare parameters associated with different treatments to understand which treatment is the most effective in boosting vaccine take up.

Additionally, we will examine the impacts of all treatments compared to the control group. To do so, we will pool individuals in the treatment groups (*INFO, INFOACCESS,* and *INFOLEADERS*) and regress the outcomes on the pooled treatment indicator.

Missing data and attrition. We will construct Lee bounds (Lee, 2009) for estimations in which the outcomes are missing for at least 10 % of the sample. To test whether attrition is systematically correlated with treatments, we will compare baseline characteristics of participants that dropped out to that of participants that stay in the study. We will also test if the attrition rate differs between treatment and control groups.

Multiple hypotheses testing correction. Because we consider multiple primary outcomes, such as vaccine uptake and intention, we will follow Anderson (2008) to correct the multiple hypothesis testing problem.

4.2. Heterogeneity analysis

We explore the heterogeneous treatment effects of several subgroups of interests to uncover potential channels. We highlight some key heterogeneities.[9]

Socioeconomic status. We anticipate that the effects of the intervention might vary by socioeconomic status, but we do not have a clear prediction on the direction—it might be either positive or negative. Socioeconomic status comprises of several variables: indicators for higher income (1 if household expenditure per capita is above the median), employment status (1 if employed), higher educational attainment (1 if education is higher than primary school or lower), and beneficiary of any social assistance.

Demographic characteristics. We anticipate that the effects of the intervention might vary by gender status and age. The positive effects are likely to be more pronounced among males than females, as observed in a recent meta-analysis study (Zintel et al., 2022). However, we do not have a clear prediction on whether the older participants aged 60 and above are more likely to take up vaccines than the younger participants.

Beliefs and knowledge about COVID-19 and COVID-19 vaccines. We anticipate that the effects of our intervention might vary by baseline beliefs, but we do not have a clear prediction on the direction. A comprehensive review study has documented mixed results on the effects of health belief model on health behaviors (Carpenter, 2010).

Formally, we estimate the following regression specification:

$$Y_{i,t=1} = \alpha + \sum_{j=1}^{3} \beta_j Tj_i + H_i \times \sum_{j=1}^{3} \gamma_j Tj_i + \theta H_i + \delta Y_{i,t=0} + X'_{vi}\tau + \varepsilon_i \quad (2)$$

[9] Our data allows us to add more heterogeneities and when we do so we will apply a machine learning method to objectively select heterogeneities.

where $T1_i$ denotes *INFO*$_i$, $T2_i$ denotes *INFOACCESS*$_i$, and $T3_i$ denotes *INFOLEADERS*$_i$. H_i denotes heterogeneity (baseline values) that alternates between key heterogeneities discussed above. Our parameters of interests are γ_j which captures heterogeneity in treatment effects.

Conducting heterogeneity analyses that involve many sample splits can lead to over-rejection of the null hypotheses. To overcome this problem, we will apply honest causal forest technique (Wager and Athey, 2018).

References

Ali, Mohammad, and Ahmed Hossain. "What is the extent of COVID-19 vaccine hesitancy in Bangladesh?: A cross-sectional rapid national survey." *medRxiv* (2021).

Anderson, M. L. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training studys." *Journal of the American statistical Association*, 103(484) :1481–1495 (2008).

Arce, Solís, J.S., Warren, S.S., Meriggi, N.F. *et al.* "COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries." *Nature Medicine* 27, 1385–1394 (2021).

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. "Using gossips to spread information: Theory and evidence from two randomized controlled trials." The Review of Economic Studies 86, no. 6 (2019): 2453-2490.

Banerjee, Abhijit, Marcella Alsan, Emily Breza, Arun G. Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin A. Olken. *Messages on COVID-19 prevention in India increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities*. No. w27496. National Bureau of Economic Research, 2020.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies* 81, no. 2 (2014): 608-650.

Carpenter, C. J. (2010). A meta-analysis of the effectiveness of health belief model variables in predicting behavior. Health communication, 25(8):661–669.

Campos-Mercade, Pol, Armando N. Meier, Florian H. Schneider, Stephan Meier, Devin Pope, and Erik Wengström. "Monetary incentives increase COVID-19 vaccinations." Science 374, no. 6569 (2021a): 879-882.

Chang, Tom, Mireille Jacobson, Manisha Shah, Rajiv Pramanik, and Samir B. Shah. "Financial Incentives and Other Nudges Do Not Increase COVID-19 Vaccinations among the Vaccine Hesitant." No. w29403. National Bureau of Economic Research, (2021).

Hossain, Mohammad Bellal, Md Zakiul Alam, Md Syful Islam, Shafayat Sultan, Md Mahir Faysal, Sharmin Rima, Md Anwer Hossain, and Abdullah Al Mamun. "COVID-19 Vaccine Hesitancy among the Adult Population in Bangladesh: A Nationally Representative Cross-sectional Survey." *medRxiv* (2021).

Khan, Hibah, Ms Era Dabla-Norris, Frederico Lima, and Alexandre Sollaci. *Who doesn't want to be vaccinated? Determinants of vaccine hesitancy during COVID-19*. International Monetary Fund, 2021.

Kim, Hyuncheol Bryant, Beliyou Haile, and Taewha Lee. "Promotion and persistence of HIV testing and HIV/AIDS knowledge: evidence from a randomized controlled trial in Ethiopia." *Health economics* 26, no. 11 (2017): 1394-1411.

Klüver, Heike, Felix Hartmann, Macartan Humphreys, Ferdinand Geissler, Johannes Giesecke. "Incentives can spur COVID-19 vaccination uptake." Proceedings of the National Academy of Sciences (2021)

Lee, D. S. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies*, 76(3) (2009):1071–1102.

Reza, Hasan Mahmud, Vaishnavi Agarwal, Farhana Sultana, Razmin Bari, and Ahmed Mushfiq Mobarak. "Why Are Vaccination Rates Lower in Low and Middle

Income Countries, and What Can We Do about It?” *BMJ* 378 (July 13, 2022): e069506. https://doi.org/10.1136/bmj-2021-069506.

Rosenstock, I. M. “Why people use health services.” *The Milbank Memorial Fund Quarterly* (1966).

Sato, Ryoko, and Yoshito Takasaki. "Vaccine Hesitancy and Refusal: Behavioral Evidence from Rural Northern Nigeria." *Vaccines* 9, no. 9 (2021): 1023.

Siddique, Abu, Tabassum Rahman, Debayan Pakrashi, Firoz Ahmed, Asad Islam. “Raising COVID-19 awareness in rural communities: A randomized experiment in Bangladesh and India.” (2020) Munich Papers in Political Economy 9/2020. Technical University of Munich.

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

Watson, Oliver J., Gregory Barnsley, Jaspreet Toor, Alexandra B. Hogan, Peter Winskill, and Azra C. Ghani. "Global impact of the first year of COVID-19 vaccination: a mathematical modelling study." *The Lancet Infectious Diseases* (2022).

Zintel, S., Flock, C., Arbogast, A. L., Forster, A., von Wagner, C., and Sieverding, M. (2022). Gender differences in the intention to get vaccinated against covid-19: a systematic review and meta-analysis. Journal of Public Health, pages 1–25.

Tables and Figures

Figure 1. Research Design

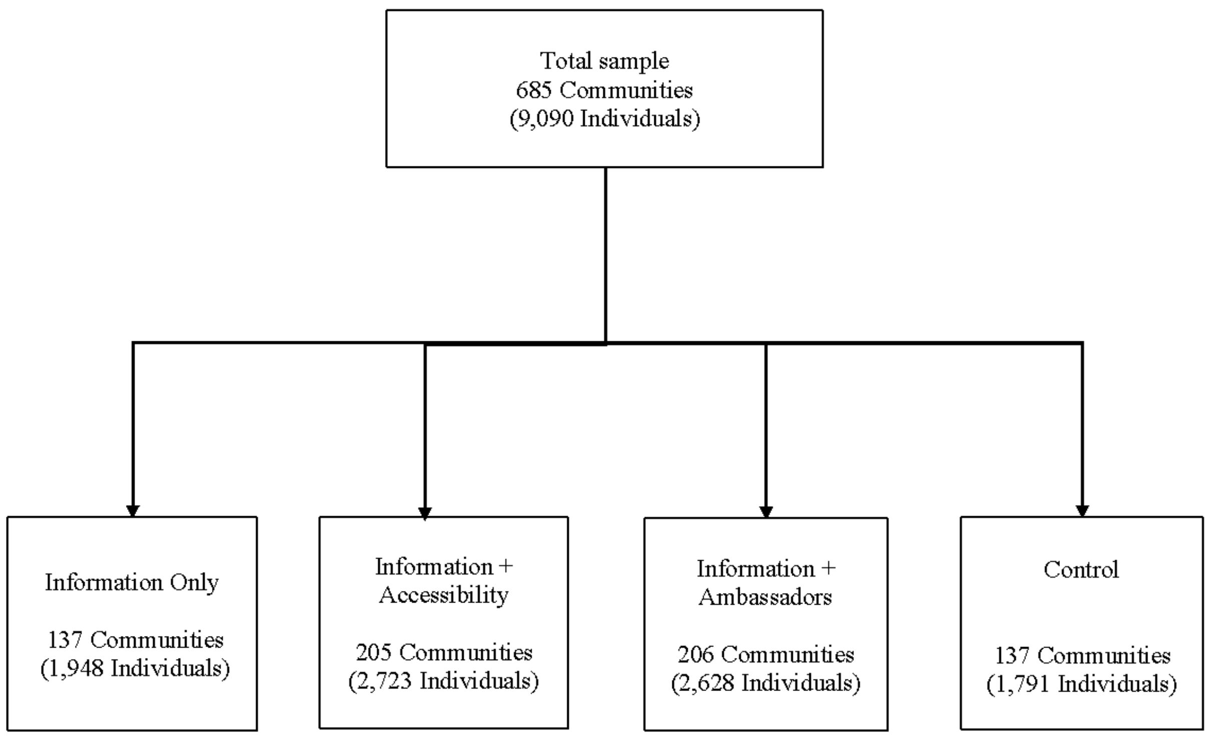

Table 1. Baseline Summary Statistics and Balance Tests

	N	Mean (pooled)	Difference between Groups (p-value)					
			Control = Information only	Control = Info+Ambassador	Control = Info+Accessibility	Info only = Info+Ambassador	Info only = Info+Accessibility	Info+Ambassador = Info+Accessibility
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Individual characteristics								
Male	9,090	0.30	0.813	0.188	0.111	0.273	0.165	0.736
Married	9,090	0.86	0.212	0.268	0.121	0.817	0.799	0.596
Muslim	9,090	0.89	0.777	0.348	0.791	0.502	0.554	0.175
Rural	9,090	0.97	0.732	0.529	0.714	0.836	0.524	0.329
Joint family	9,090	0.43	0.977	0.793	0.479	0.758	0.436	0.615
Own house	9,090	0.97	0.564	0.795	0.903	0.308	0.597	0.631
Received government assistance	9,090	0.12	0.538	0.207	0.059	0.526	0.187	0.430
Age	9,090	27.58	0.423	0.366	0.052	0.946	0.241	0.235
Have secondary education	9,090	0.31	0.123	0.183	0.203	0.715	0.635	0.911
Employed	9,090	0.22	0.979	0.697	0.960	0.707	0.934	0.597
High monthly income household	9,090	0.93	0.415	0.757	0.517	0.591	0.840	0.725
Follow COVID-19 protocols (index)	9,090	1.18	0.364	0.170	0.829	0.705	0.238	0.086
Joint-Test Prob > F			0.799	0.559	0.259	0.898	0.736	0.548
Panel B: Village characteristics								
Proportion of muslims	685	0.75	0.314	0.181	0.412	0.816	0.779	0.563
Nearest distance to COVID-19 vaccine centers (in km)	685	2.69	0.181	0.385	0.465	0.546	0.407	0.836
Nearest distance to community clinic (in km)	685	2.33	0.916	0.985	0.941	0.919	0.849	0.911
Nearest distance to railway station (in km)	685	63.47	0.780	0.503	0.356	0.704	0.508	0.730
Nearest distance to secondary school (in km)	685	3.95	0.857	0.746	0.742	0.908	0.907	0.998
Nearest distance to college (in km)	685	6.11	0.694	0.539	0.700	0.872	0.938	0.772
Nearest distance to post office (in km)	685	2.13	0.330	0.808	0.864	0.232	0.402	0.670
Nearest distance to bank (in km)	685	4.37	0.656	0.605	0.659	0.957	0.317	0.259
Nearest distance to police station (in km)	685	10.36	0.993	0.979	0.593	0.987	0.594	0.560
Nearest distance to hospital/doctor (in km)	685	2.79	0.979	0.746	0.730	0.776	0.758	0.963
Proportion of poor families	685	0.27	0.051	0.526	0.159	0.126	0.471	0.374
Proportion of landless households	685	0.22	0.281	0.537	0.536	0.545	0.556	0.992
Village head lives in the village	685	0.93	0.821	0.859	0.733	0.673	0.559	0.853
Number of families in the village	685	730.94	0.987	0.957	0.588	0.971	0.581	0.553
Joint-Test Prob > F			0.607	0.974	0.950	0.950	0.958	0.993

Note: Variable *Follow COVID-19 protocols (index)* has the maximum value of 8. Column 1 reports total number of observations. Column 2 reports average value of each variable for the whole sample (pooled). Columns 3 to 8 report *p*-values of the coefficient from regressing each baseline variable on treatment group indicators. Robust standard errors are clustered at village level. Joint Orthogonality Test Prob > F refers to the p-value of F-test of a regression of treatment indicators being compared on all baseline variables (separately for individual and village characteristics) reported in this table. This test provides an overall evaluation of the balance between groups across all baseline variables. *, **, and *** denote statistical significance at 10 %, 5 %, and 1 % levels, respectively.