Pre-Registration of Ghana Financial Incentives Trial Wave II: Spillover and Tuberculosis Screening

Raymond Duch [*] Philip Clarke [†] Edward Asiedu [‡] Thomas S. Robinson[§]
Ryota Nakamura [¶] Thomas Rouyard [‖] Carlos Yevenes [**] Adrian Barnett [††]
Laurence Roope [‡‡] Mara Violato [§§]

February 18, 2024

Abstract

The Wave I Ghana Financial Incentive trial confirmed that financial incentives have a positive effect on COVID-19 vaccine intentions, reported vaccination status and verified vaccination status. Wave II provides an opportunity to understand whether this financial incentive effect generalizes to other types of health behavior: Does a similar financial incentive design with tuberculosis screening produce similar positive results? The Wave I trial was designed to identify any spillover effects across treatment arms or on untreated individuals within treated communities. We found no evidence of negative spillover effects of cash incentives on individuals who received no cash compensations. An unanswered question is whether the financial incentives affect within-subject behavior. Wave II examines the willingness of those who received financial incentives in the initial TB screening treatment to adopt similar preventative health behaviors six months later. Wave II is also designed to evaluate the relative impact of financial incentives compared to simple informational reminders.

[*]Nuffield College, University of Oxford. raymond.duch@nuffield.ox.ac.uk.
[†]University of Oxford and University of Melbourne. philip.clarke@ndph.ox.ac.uk
[‡]University of Ghana. edasiedu@ug.edu.gh
[§]Department of Methodology, London School of Economics and Political Science. t.robinson7@lse.ac.uk
[¶]Hitotsubashi Institute for Advanced Study. ryota.nakamura@r.hit-u.ac.jp
[‖]Hitotsubashi Institute for Advanced Study. thomas.rouyard@r.hit-u.ac.jp
[**]University of Santiago Chile. carlos.yevenes@usach.cl
[††]Queensland University of Technology. adrian.barnett@qut.edu.au
[‡‡]University of Oxford. laurence.roope@dph.ox.ac.uk
[§§]University of Oxford. mara.violato@dph.ox.ac.uk

Motivation

There is a large body of experimental evidence suggesting that financial incentives can promote certain health care behaviors (Lagarde, Haines and Palmer, 2009; Giles et al., 2014). With respect to Lower and Middle Income Countries (LMICs), there have been some encouraging results from studies regarding incentives and vaccination uptake (Gibson et al., 2017). The COVID-19 pandemic highlighted the importance of understanding whether financial incentives are an effective policy tool for promoting vaccinations (Brewer et al., 2022), particularly in the African context. A number of studies of financial incentives for COVID-19 vaccines have been conducted, primarily in High Income Countries (HICs). The recent Mardi et al. (2022) systematic review of 25 recently published studies of incentives for COVID-19 vaccinations concluded that high financial incentives are attributed to higher vaccination rates while low amounts of financial incentives have small or non-significant effects. The impact of cash incentives though have been understudied in LMICs (Merriam and Behrendt, 2020) which is unfortunate since cash incentives have been considered as a means of improving the efficiency and equity of, for example, the roll-out of COVID-19 vaccinations in Africa (Arezki, 2021).

To evaluate whether cash incentives affect the willingness to get COVID-19 vaccines, we implemented the Ghana Financial Incentive Wave I randomized control trial in consultation with the University of Ghana and the Ghana Health Service (Duch et al., 2023). The experiment evaluated the impact of cash incentives on vaccine uptake. Results from this two-stage randomized cluster trial in rural Ghana indicated that financial incentives increased COVID-19 vaccine uptake (Duch et al., 2023). We first measured intentions because an expressed willingness to get vaccinated is correlated with actual vaccination outcomes (Athey et al., 2022; Moehring et al., 2021). We observed that non-vaccinated participants, assigned to the financial incentive treatment arms, are about 10% more likely to express an intention to get the COVID-19 vaccine. Participants in a Cash treatment arm had an average vaccination intention rate of about 81% compared to 71% for those in the Placebo treatment arm. Two-months later, post-treatment, we collected self-reported vaccination rates. The Low Cash arm ($3.00) reported the highest vaccination rate of 41.8% compared to 36.5% for the Placebo arm (difference from Placebo: 5.3; 95% CI: 1.6, 8.2; P = 0.003). The combined Cash treatment arm ($3.00 and $10.00) had a vaccination rate of 40.0% (difference from

Placebo: 3.6%; 95% CI: 0.001, 6.9; P = 0.03). The trial was also designed to measure spillover effects of the financial incentives to participants in the Placebo and Health message treatment arms as well as to non-treated participants in villages treated with the financial incentives. We reported no spillover effects to either non-treated proximate individuals or participants treated in the non-financial treatment arms.

In this pre-registration we describe Wave II of the Ghana Financial Incentive project. Our goal in Wave II is to explore important findings from the Wave I Ghana Financial Incentive trial. The Wave I trial indicated that financial incentives have a positive effect on COVID-19 vaccine intentions and reported vaccination status. Wave II provides an opportunity to understand whether this effect generalizes to other types of health behavior: Does a similar financial incentive design with health screenings produce similar positive results? A key finding in Wave I is that a $3.00 cash payment had a larger treatment effect than $10.00. We retain the $3.00 cash incentive in the Wave II implementation.

There is extensive evidence that non-pecuniary "nudges" might be sufficient to generate the treatment effects we observed in Wave I – possibly even larger (Reñosa et al., 2021; Dai et al., 2021). These are, of course, much cheaper and easier to scale up (Milkman et al., 2022; Sasaki, Saito and Ohtake, 2022). Wave II incorporates a telephone/text nudge treatment that reminds participants that the screening clinic is available in their village.

The spillover effects of financial incentives for healthcare behaviours, such as vaccinations and disease screening, is an important policy concern. The Wave I trial was designed to identify any spillover effects across treatment arms or on untreated individuals within treated communities. We found no evidence of negative spillover effects of cash incentives on individuals who received no cash compensations. This is consistent with the recent findings of Schneider et al. (2023). For Wave II we extend the design to explicitly address three different forms of spillover (Benjamin-Chung et al., 2018). First, we measure the health screening behavior of other household members – on the assumption that there is within household spillover (Benjamin-Chung et al., 2017; Sinclair, McConnell and Green, 2012). Second, as we did in Wave I, our design allows us to estimate spillover to proximate neighbors. Thirdly, an unanswered question from Wave I is whether the financial incentives affect within-subject behavior – the willingness of those who received financial incentives to adopt similar preventative health behaviors in the future. Wave II is designed to address within-

subject spillover effects from financial incentives. In Wave II we will observe compliance with a different health screening protocol (a hypertension screening protocol) 6 months after the initial intervention.

Access to vaccination facilities has been shown to play an important role in compliance with COVID-19 vaccine campaigns (Mobarak et al., 2022). The Wave II trial controls for the causal effect of access on Health Screening. In the trial all participants will have identical access to Health Screening because we will be implementing screening clinics in their villages.

There are two health screening information treatments administered as part of the trial. The initial treatment phase includes video treatment arms that highlight the importance of tuberculosis screenings. Tuberculosis is responsible for 4.91% of deaths in Ghana – the public health challenge is detection with only 34% of cases estimated to be detected (Kuupiel et al., 2023; Institute for Health Metrics and Evaluations, 2018). A second post-treatment phase treats subjects with an information video that highlights the importance of hypertension screening. A recent meta-study of hypertension in Ghana estimated that 27% of the adult population suffered from the disease; two-thirds of hypertension cases in Ghana are undiagnosed; and hypertension is the leading cause of hospital admissions and deaths in Ghana (Bosu and Bosu, 2021; Ghana Health Service, 2018).

Randomized Control Trial

The trial has a factorial design with five treatments that is summarized in Table 1:

- T1: Placebo video discussing solar charging devices
- T2: 2 X 2 Factorial Encouragement Treatments
  - T2a: Health Information Message video encouraging health screening with NO text/voice reminders and NO cash incentive.
  - T2b: Health Information Message video with a text/voice reminder and NO cash incentive.
  - T2c: Health Information Message video with NO text/voice reminder and with cash incentive.

– T2d: Health Information Message video with text/voice reminder and with cash incentive.

			Health Information Video	
			NO Text Reminder	**YES Text Reminder**
Placebo	**Health Information Video**	**NO Cash Incentive**	T2a: Health Information; NO text/voice reminders; NO Cash	T2b: Health Information; text/voice reminder; NO cash
T1: Placebo video discussing solar charging devices		**YES cash Incentive**	T2c: Health Information; NO text/voice reminder; with cash	T2d: Health Information; text/voice reminder; with cash

Table 1: Research Design

Hypotheses The goal of the trial is to determine how to increase compliance with Health Screenings. We test for the impact of three encouragement treatments on compliance with Health Screenings: a simple Health Information video; a $3.00 cash incentive; and text/voice reminders. The control treatment is a placebo video.

- H1-H4: Health Screening rates by subjects in the Health Information treatment arm, in the $3.00 (40 Ghana Cedis) Cash treatment arm, and in the text/voice reminder arm will be higher than for those in the Placebo arm.
  - H1: T2a > T1
  - H2: T2b > T1
  - H3: T2c > T1
  - H4: T2d > T1
- H5-H6: Subjects in the text/voice reminder treatment arms will have higher Health Screening than subjects with only the Health Information video treatment.
  - H5: T2b > T2a
  - H6: T2d > T2a
- H6-H7: Subjects in the $3.00 cash incentive treatment arm will have higher Health Screening than subjects with only the Health Information video treatment.
  - H6: T2c > T2a
  - H7: T2d > T2a
- H8: Subjects in the $3.00 cash incentive treatment arm will have higher Health Screening than subjects in the text/voice reminder treatment arms.
  - H8: T2c > T2b
- H9: Subjects in the $3.00 cash incentive plus text/voice treatment arm will have higher Health Screening than subjects in the treatment arm with only a text/voice reminder.
  - H9: T2d > T2b

- H10: Subjects in a $3.00 cash incentive treatment arm will have higher Health Screening than subjects in no cash arms.
  - H10: T2c + T2d > T2a + T2b
- H11: Subjects in a voice/text reminder treatment arm will have higher Health Screening than subjects in no voice/text treatment arms.
  - H11: T2b + T2d > T2a + T2c

We are proposing to test a number of hypotheses and are aware of the ongoing debate regarding the reporting of statistical significance (Perneger, 1998). We will adopt a conservative and transparent strategy in reporting results along with their statistical significance.

Spillover Effects Another aim of the study is to assess whether our treatment arms have spillover effects. In the previous Wave I of the Ghana trial, we examined whether there was spillover to subjects who were proximate to the financial incentive treatment, i.e., who were either in the non-financial treatment arms or who were not treated.

In this Wave II of the Ghana trial we examine three types of spillover effects: household spillover; geographic proximity spillover; and within-subject spillover.

- H12: Within each of the sampled households we ask enumerators to itemize all the other members of the household. Their compliance with the Health Screenings is also measured post-treatment.
  - H1 through H4 will hold for non-treated household members.
- H13: Within each of the sampled village clusters that are assigned to treatment arms T2a through T2d, we assign 25% of the sample to the placebo treatment video.
  - H1 through H4 will hold for non-treated subjects in village clusters assigned to treatment arms T2a through T2d.
- H14: In addition to verifying the screening status of treated subjects we will have a complete record of screening compliance by all members of all the village cluster. The count of untreated screened subjects will be used to estimate spillover effects.

  - Our primary hypothesis is that untreated subjects in the TB cash treatments (T2c and T2d) will have lower screening rates than untreated subjects in the no cash treatments (T2a and T2b) and lower than those original treated with the placebo (T1).

- H15: Six months following the initial intervention we re-contact all subjects and ask them to participate in an un-related hypertension screening.

  - Our primary hypothesis is that subjects in the TB cash treatments (T2c and T2d) will have lower hypertension screening rates than those in the original no cash treatments (T2a and T2b) and lower than those original treated with the placebo (T1).

Design

The randomized control trial is designed to measure 1) the direct impact of financial incentives on TB health screening rates in the Ghana rural population and 2) three forms of spillover effects of financial incentives on TB screening: within household; proximate neighbors; and within-subject. The CONSORT diagram in Figure 1 summarizes the overall design of the RCT. We will sample a total of 9,000 subjects for the randomized control trial in 450 villages from 6 rural Ghana Districts. Subjects will be randomly assigned to one of the five treatment arms presented in Table 1.

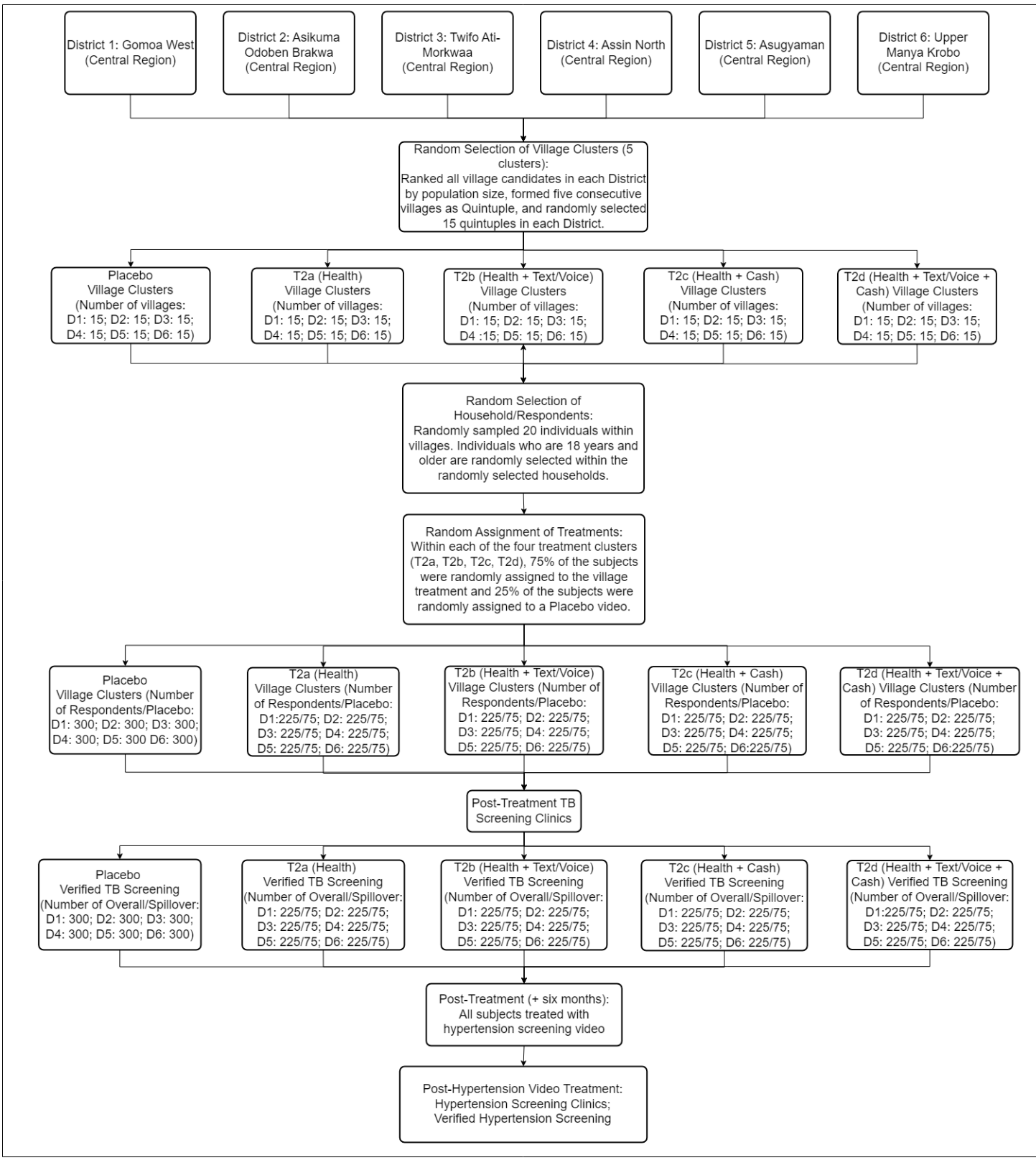

Figure 1: CONSORT diagram for the Ghana Financial Incentives Trial II. Phase I: TB Video Treatment. Phase II: a) Post-treatment TB screening clinic. Phase III: Hypertension Video Treatment. Phase IV: Hypertension screening clinics. D1 through D6 denotes Districts.

In Phase I of the trial participants will be contacted by enumerators and will be treated with one of the five treatment videos. They will also receive an invitation to get a TB screening the following week in the village on an assigned two day period. Enumerators will leave a pamphlet with the details and timing with all participants. Tuberculous screening clinics are organized in each of the treated villages as part of Phase II. A post-treatment hypertension screening video is administered to all participants six months after the Phase I intervention – participants are encouraged to get a hypertension screening in village clinics. Finally in Phase IV of the trial, two-day hypertension screening clinics are organized in each of the treated village.

Treatments

Video Components The experiment has three treatment videos:

- T1: Placebo: A 45-second placebo video that provides general information about the benefit of using solar power to charge household electrical appliances.
- Treatment 2a and 2b: A 45-second standard health TB promotional and information video (modeled on the videos produced by English National Health Service).
- Treatment 2c and 2d: Cash Incentive treatment – the first 45 seconds are identical to the Health video – the last 15 seconds inform viewers that that they will earn $3.00 if they receive the TB screening administered in their village during a two-day period the following week.

We have developed the draft storyboard for the baseline health treatment video:

- Scene 1: Introduction (5 seconds)
  - Open with an overhead shot of a bustling city
  - Fade to a close-up of a person coughing and looking unwell
  - Cut to text on the screen: "Tuberculosis: A Hidden Threat"
- Scene 2: Explanation of TB (15 seconds)
  - Cut to an animated graphic showing a bacterium causing TB

    - Narrator starts explaining what TB is and how it spreads
    - Show visual aids such as illustrations or animations to explain the symptoms, such as coughing, chest pain, fatigue, and weight loss
    - Highlight the fact that TB is spread through the air when an infected person coughs or sneezes
  - Scene 3: Importance of Diagnosis and Treatment (10 seconds)
    - Show an illustration of a person going to a doctor for a check-up
    - Narrator explains the importance of early diagnosis and treatment for TB
    - Highlight that TB can be treated and cured with the right medication and care
  - Scene 4: Preventing TB (10 seconds)
    - Cut to an illustration of a person covering their mouth while coughing
    - Narrator explains the steps to prevent TB, such as covering the mouth while coughing, frequent handwashing, and avoiding close contact with infected individuals
    - Highlight the importance of seeking medical attention if one experiences symptoms of TB
  - Scene 5: Closing (5 seconds)
    - Cut to text on the screen: "Protect Yourself and Your Community"
    - Show a graphic of a community coming together to fight TB
    - End with the logo of the organization promoting the TB health message Note: The length of each scene and the visuals can be adjusted based on the desired emphasis and pacing of the message.

Text/Voice Messaging Component A second component of the treatment protocols is a text and voice messaging reminder that is sent out to subjects, in Treatments 2b and 2d. They receive the voice/text message three days prior to when the TB screening clinics are scheduled for their villages and also on the day of the screening. Enumerators contact subjects in this treatment arm by phone. If the subject is contacted they receive a scripted reminder of the TB screening clinic in their village. If not contacted, they receive both a reminder text

along with a short recorded voice reminder. We provide both forms of reminders since some of the subjects will be illiterate. Based on our previous trial with the identical population all participants we expect all participants to have access to a cell phone. In many cases this will be the cell phone of an immediate family member or neighbor.

There have been a number of recent randomized control trials evaluating the effect of reminder messages on vaccination rates. The recent large scale trial by Milkman et al. (2022) assessed the impact of 22 different reminder messaging protocols on pharmacy flu vaccination rates. The most effective text reminder was worded as follows: "A flu shot is waiting for you at Walmart." We have adopted a variant on this effective messaging strategy: "We are calling to remind you that a Tuberculosis screening by District Health Officials will be waiting for you on [day/date] at your [name] village square."

Sampling Strategy

We will contact 9,000 sampled individuals in the 450 villages and 6 Ghana Districts. Enumerators will be instructed to randomly select households using a random walk procedure (Bennett et al., 1991). Figure 2 summarizes the sampling stages that we will implement in the experiment.

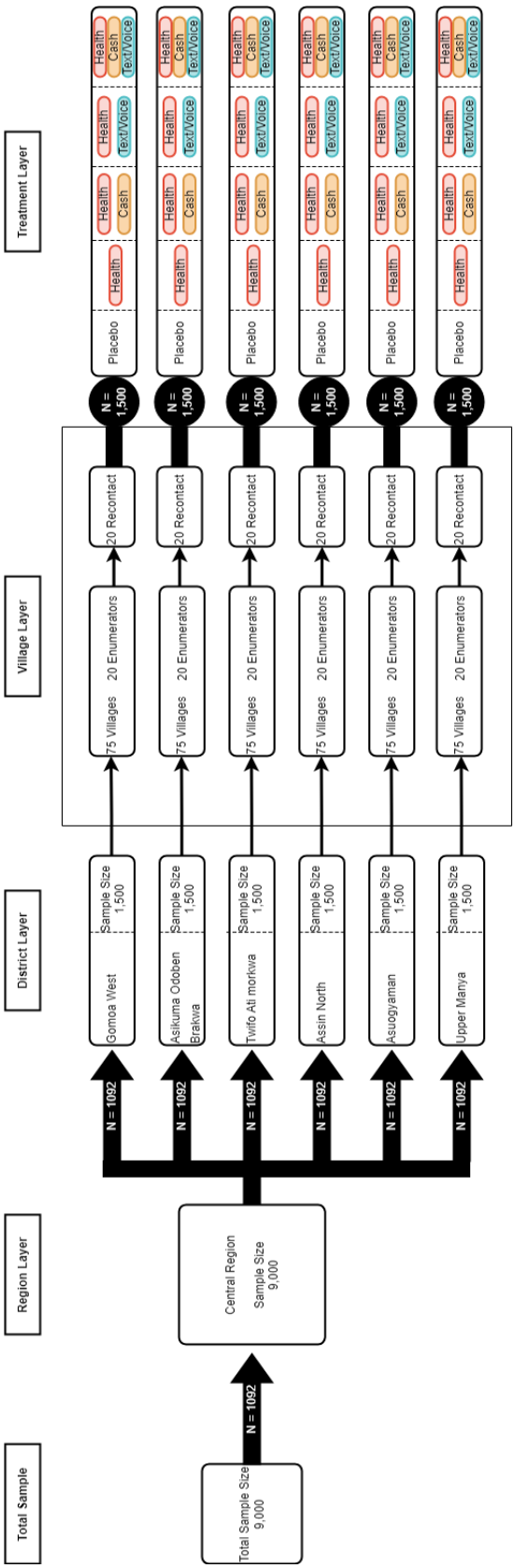

Figure 2: Experiment Flow Chart

The information treatments have been designed to be delivered door-to-door, by enumerators working with the University of Ghana. By adopting the door-to-door implementation we will ensure a high degree of control over which households receive each of the informational treatments and which do not. For predominantly rural Ghana villages, door-to-door visits are also the most feasible choice given the low levels of literacy.

Trial design, eligibility, randomization and recruitment. There are 16 regions in Ghana. Districts are classified in three types: Ordinary Districts with a minimum population of seventy-five thousand (75,000) people, Municipal Districts with a minimum population of ninety-five thousand (95,000) people, and Metropolitan Districts with a minimum population of one hundred and fifty thousand(150,000) people.

Working with the District offices of the Ministry of Health we selected one region (Central) and six districts:

- District 1: Gomoa West – Central Region
- District 2: Asikuma Odoben Brakwa – Central Region
- District 3: Twifo Ati-Morkwaa – Central Region
- District 4: Assin North – Central Region
- District 5: Asugyaman – Central Region
- District 6: Upper Manya Krobo – Central Region

This is a cluster randomized trial – villages are randomly assigned to one of the treatment arms. The CONSORT diagram (Figure 1) described the design and the allocation of clusters to each arm.

Working with our partners in each district, we will identify the list of District villages that could feasibly be enumerated by our enumeration team (the primary consideration here was either road access or the quality of the road access). Then, within each District, we will generate a complete list of all of the villages that are candidates for enumeration. The villages were then ranked according to their population size (population statistics provide by the 2020 Ghana census). We then formed groups of five villages by putting five consecutive

villages on these lists in the same quintuple. In a typical district we would have approximately 50 quintuples.

In each district we randomly selected 15 quintuples with probabilities weighted by the quintuple's share of the total population of the villages being considered in the district. This initial sample of quintuples was then adjusted in consultation with the District Health officials. The adjusted selection criteria are in part driven by cost considerations – for budget reasons we were constrained to ensure reasonable travel distances between the five village clusters. We also need to ensure, in collaboration with our District Health officials, the feasibility of providing TB health screening clinics in each village. Within each of the chosen quintuples, we randomly selected a village to be assigned one of the five treatment arms described in Table 1. This will result in a total sample of 450 villages from the 6 different districts. Within each of the four treatment villages (T2a, T2b, T2c, and T2d), 75% of the subjects were randomly assigned to the village treatment and 25% of the subjects were randomly assigned to a Placebo video.

Enumerator IDs will be assigned to village clusters. At the beginning of each enumeration day, enumerators download their Qualtrics questionnaire assignments with the appropriate embedded videos. At the end of the day these completed questionnaires will be uploaded to the Candour server. No survey data are left resident on the individual tablets. Enumerators will be provided with a random walk protocol for selecting households in each village.

Within each of the 450 villages we will randomly sample 20 individuals. Households were randomly selected and then within households we randomly selected individuals 18 years and older. There are five treatment arms with 90 clusters in each treatment arm.

Power Calculations

In the Wave II Ghana Incentive Trial, we will sample 9,000 individuals in 75 villages and 6 Ghana Districts. Village clusters are randomly assigned to one of the five treatment arms described in Table 1.

We run power calculations by drawing outcomes from a binomial distribution:

$$Y_{ij} \sim \text{Binom.}(n = 1, p = P(\text{TB Test})_{ij}), \tag{1}$$

where the underlying probability of getting a tuberculosis test is modelled linearly using the identity link function as,

$$P(\text{TB Test})_{ij} = \beta_0 + \beta_{2a}\text{T2a}_{ij} + \beta_{2b}\text{T2b}_{ij} + \beta_{2c}\text{T2c}_{ij} + \beta_{2d}\text{T2d}_{ij} + v_j. \tag{2}$$

where:

- i and j index individuals and villages respectively
- $P(\text{TB Test})_{ij} = 1$ if the linear components sum to greater than 1 and $P(\text{TB Test})_{ij} = 0$ when the components sum to less than 0
- β_0 is the placebo probability of getting tested for tuberculosis in the week following the intervention
- β_{2a} is the treatment effects of watching the Health message about tuberculosis testing with NO cash incentive and No text/voice reminders.
- β_{2b} is the treatment effects of watching the Health message about tuberculosis testing with NO cash incentive and with text/voice reminders.
- β_{2c} is the treatment effects of watching the Health message about tuberculosis testing with a cash incentive and with NO text/voice reminders.
- β_{2d} is the treatment effects of watching the Health message about tuberculosis testing with a cash incentive and with text/voice reminders.
- v_j are village-level fixed effects with mean 0, and standard deviation set to induce a 0.015 intraclass correlation coefficient (ICC) for the binary outcome (which mirrors the observed ICC in the Wave I study).

We noted our hypotheses earlier during our discussion of Table 1. We focus our attention on the coefficient for the Health Message only treatment (β_{2a}) as this is expected to be smaller,

and thus harder to statistically distinguish from the placebo group.[1] To recover an estimate of the minimum effect size for the Health only treatment effect ($T2a$), with 0.8 power at the 95% confidence level, we begin by assuming $\beta_{2a} = 0.06$. We then generate 1000 hypothetical datasets, each with 9000 subjects across 450 treatment villages. We assume equal probability of assignment of villages (and the subjects within them) to the five treatment arms, and within each non-placebo assigned village we assign 5 of the 20 subjects to the placebo treatment. For each simulated dataset, we realise the outcome for each subject, and then fit the linear model in Equation 2.

We estimate the power of the simulated experiment as the proportion of Health coefficients ($\beta_2 a$) across these 1000 models with p-values below 0.05 (using the conventional two-sided test). If this power is above 0.8, we decrease the hypothesised effect size by 5% and repeat the simulation, until the estimated power falls below 0.8. The final well-powered effect size is our estimate of the minimum effect size we can reliably detect.

Figure 3 reports these minimum effect size estimates, under increasing baseline probabilities of getting tested (the placebo). If the baseline probability of vaccination is 0.5 we expect to be well-powered to observe an effect size of 0.05 and above. As the baseline probability increases, so does the precision of our estimates. At a baseline probability of 0.9, we are powered to detect effect sizes greater than 0.026. Note that, by symmetry, our power estimates are the same if we were to decrease the baseline probability towards 0. Therefore if the baseline vaccination rate is 0.1, we would also expect to be able to distinguish effect sizes greater than 0.026.

The simulations are conducted in `R` and the simulation code is available in the Appendix.

[1] To simplify our analysis, we assume the effect $T2b$ is two times, $T2c$ three times, and $T2d$ four times, the size of $T2a$.

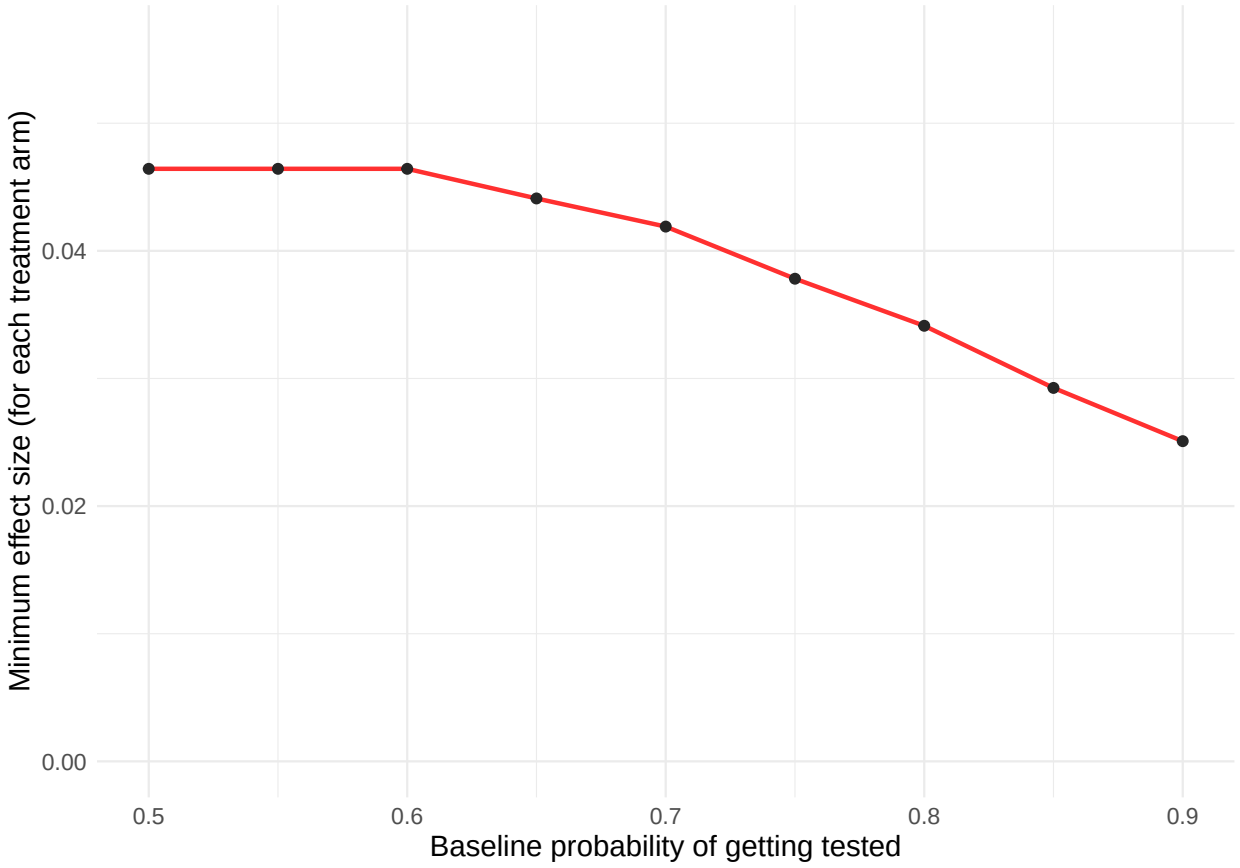

Figure 3: Estimated minimum effect sizes with 0.8 power and $\alpha = 0.05$, with $N = 7000$

Data Collection Stages

Figure 4 summarizes the phases of data collection that will make up the Wave II Ghana Financial Incentive trial. We expect data collection to begin on April 1, 2024 and to be completed by June 1, 2024.

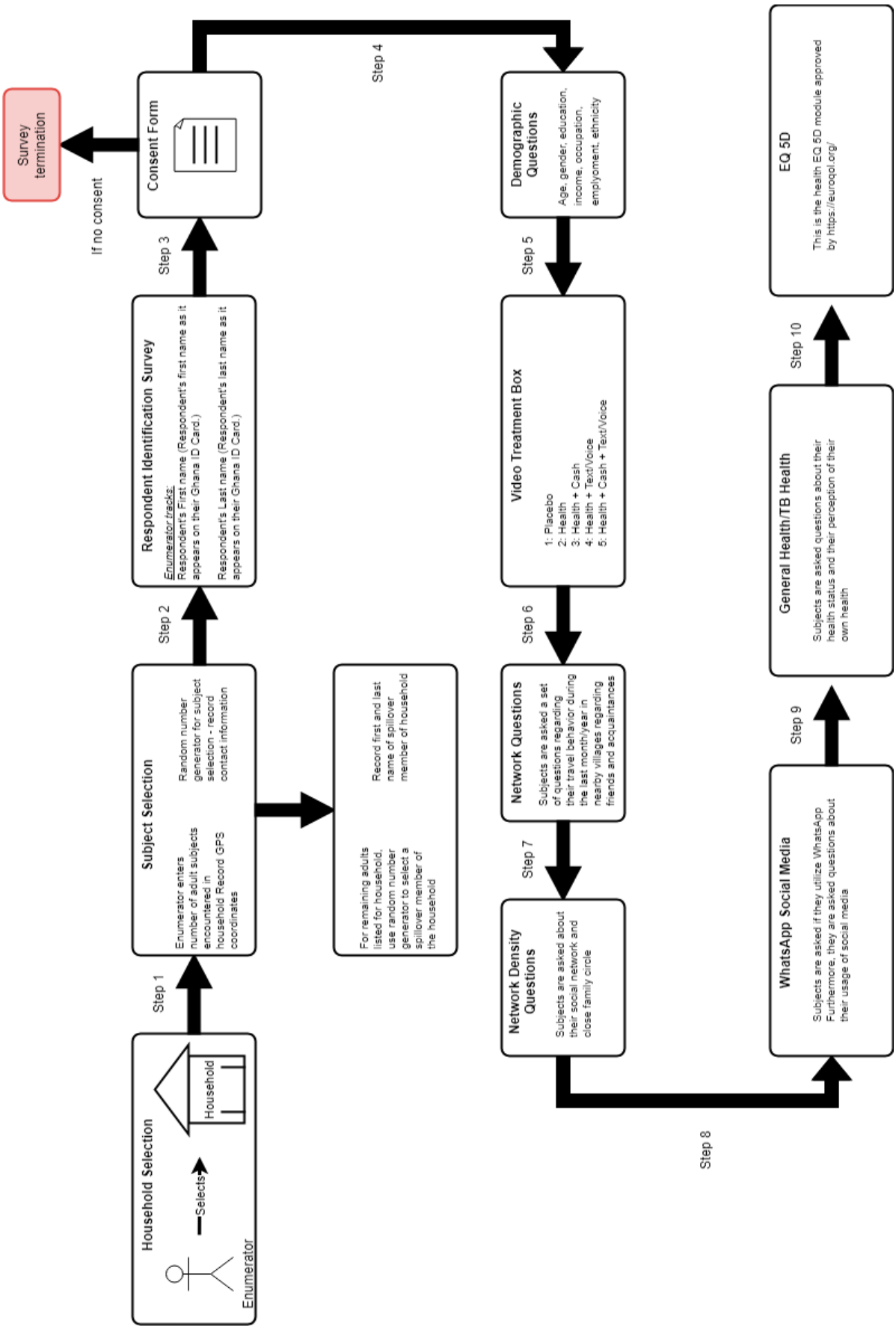

Figure 4: Flow Chart of Data Collection Phases

Phase 1: Treatment

The initial data collection and treatment phase of the experiment will begin on April 1, 2024 and continue through to approximately June, 2024. As described earlier the village clusters are randomly assigned to one of five treatment arms. And within each village we selected a supplemental sample of 5 subjects.

Enumerators are instructed to share the Survey Tablet with the subjects. Instructions and questions are read aloud to the subjects. The enumerators will enter the subjects' responses into the Tablet application. The phase one intervention consists of the following elements:

- Enumerators prepare a complete itemization of adult members of the household. With a random number generator they select one adult to participate in the study. They record the name of the study participant. From the remaining adults in the household, they randomly select a within-household spill-over individual and record their full name as part of the survey (Step 2).
- Participants are presented with the purposes of the study and an informed consent form that they must agree to before proceeding to the survey (Step 3).
- This is followed by a set of demographic questions including age, gender, education, income, occupation, employment status and ethnicity (Step 4).
- Subjects are then presented with one of the short 45-second treatment videos (Step 5).
- Subjects in the Cash Incentive treatment arms (T2c and T2d) are informed that they will receive their $3.00 TB testing compensation via cell phone if show up for the TB testing to be conducted the following week in their village.
- Subjects in both are asked a set of questions regarding their intention to get the TB testing the following week in their village.
- A set of questions are included that measure the subject's travel patterns within and outside of the village (Step 6).
- A related set of questions measure the density of the subject's social and family network (Step 7).

- Subjects are asked a battery of questions regarding their social media accounts and activity (Step 8).
- Subjects are asked a battery of questions regarding their general health including any history of Tuberculosis (Step 9).
- Subjects are asked the complete battery of EQ-5D Health Questions (Step 10).
- The survey ends with requests for contact information: cell phone number; WhatsApp; and e-mail addresses.
- Subjects are provided with a contact telephone/WhatsApp number plus e-mail address that they can contact if they have further questions.
- Subjects receive their participation compensation of $1.00 in cash.

All subjects are compensated for their participation in the experiment. Subjects earn 1.00 USD for agreeing to participate in the study. This is paid in cash after completing the first survey interview. We ask them for their permission to contact them at a later date, and inform them that they would be compensated again if they are contacted and complete the second interview.

The data collected during these visits are recorded on a Qualtrics App that is resident on each of the Tablets. At the end of the day, when the enumerators have internet access, these data are shared with the Qualtrics API and are registered on the Candour project Qualtrics account database.

Phase 1b: Text/Voice Reminders

Subjects in the text/voice reminder treatments ($T2b$ and $T2d$) will receive a text and voice reminder 3 days before, and on the day of, the scheduled screening clinic. The reminder message will be the following: "We are calling to remind you that a Tuberculosis screening by District Health Officials will be waiting for you on [day/date] at your [name] village square."

Phase 2: TB Screening Clinics

Working with each of the six Health Districts, we will organize a two-day TB screening clinic in each of the 450 villages within one-week of the treatment interventions. The TB screening phase will include the following:

- Two health personnel will be scheduled to visit each community on two days the week following the video treatment intervention.
- The local treatment clinic will be advertised to the local village via conventional health information channels. This includes broadcasting information to the village prior to the arrival of the TB health team and also broadcasting the information during the two days that the clinic takes place. The villages are relatively small and broadcasting is conducted via village loudspeakers and a mobile announcement vehicle.
- A standard TB screening questionnaire will be administered to all who show up for the clinic.
- Those who are screened as possibly having TB will receive a GeneXpert (`https://www.ultragenyx.com/`) TB test without charge.
- The health officials will maintain a list of the names of all individuals tested during the two-day clinic.
- This confidential digitized list will be shared via a secure link with the research director. The list of screened individuals will be matched to the list of trial participants including the within-household spillover individuals.
- Subjects in the $3.00 cash treatment groups who get the TB screening would be sent their payment via cell phone payment – they would be informed of this at the end of the initial intervention survey.

Phase 2b: Untreated Spillover

We are scheduling TB screening clinics for two days in each treated village. In addition to verifying the screening status of treated subjects we will have a complete record of screening

compliance by all members of the village cluster. The count of untreated screened subjects will be used to estimate spillover effects.

Phase 3: Spillover Treatment

Six months following the initial intervention we re-contact subjects and ask them to participate in a hypertension screening:

- Subjects are reminded of the purposes of the study and presented with an informed consent form that they must agree to before proceeding to the survey.
- This is followed by a set of demographic questions including age, gender, education, income, occupation, employment status and ethnicity.
- The complete battery of EQ-5D Health questions are administered as part of the 6-month follow-up.
- Subjects are then presented with a short 45-second hypertension health message video.
- Two health personnel will be scheduled to visit each community one week following the hypertension health video treatment intervention;
- The local treatment clinic will be advertised to the local village via conventional health information channels. This includes broadcasting information to the village prior to the arrival of the TB health team and also broadcasting the information during the two days that the clinic takes place. The villages are relatively small and broadcasting is conducted via village loudspeakers and a mobile announcement vehicle;
- A standard hypertension screening will be administered to all who show up for the clinic and health officials will follow their standard procedure for following up on positive test results;
- The health officials will maintain a list of the names of all individuals tested during the two-day clinic;
- This confidential digitized list will be shared via a secure link with the research director. The list of screened individuals will be matched to the list of trial participants including the within-household spillover individuals.

Data Collection Quality Assurance

The central research team will conduct regular data checks every day, outputting summary statistics for the most recent data, and regularly sharing them with the team. Quality checks include the following fields, and example reports:

- Numbers of surveys completed each day.
- Number of surveys by enumerator and community.
- Treatment assignment statistics to determine random assignment to treatment/placebo implemented properly.
- GPS tracking to verify random walk implemented properly.
- Survey length statistics, flagging any submitted survey that lasts less than 25 minutes.
- Check if any people required referrals, either to non-urgent medical care, or more urgent problems. For urgent problems we will reach out immediately to an NGO.
- Check contact numbers for formatting errors, and check for duplicates that might suggest problematic submissions.
- Descriptive statistics on how many contact numbers respondents give us, which will be used in follow-up surveys.
- Item non-response rate.
- Return intentions.
- Checking for logical inconsistencies in survey responses.

The principal investigators will be in Ghana to oversee the project. Ray Duch will directly supervise enumerator training, the survey pilot, and the beginning of data collection.

The University of Ghana team will conduct regular quality checks of submitted data and forward any concerns to the principal investigators for their attention and discussion on a continual basis. The University of Ghana team will oversee data collectors in field sites. There will be supervisors for the fieldworkers, who will escort them during all data collection and ensure that they follow the sampling technique as per the protocol. Survey supervisors

will send weekly progress reports to the research team. Progress reports include quality notes itemized by survey ID. Progress reports also include recruitment refusal rates with detailed statistics of recruitment characteristics.

The research team will build multiple quality checks into the survey instrument and check them regularly. (Early in data collection these built-in checks allow us to identify any enumerator who is submitting problematic data. Such an enumerator would be taken off the project. All their surveys would be thoroughly checked for quality and we would manage to replace all their problematic surveys with new ones). We will conduct continual confirmation of geographic sites of interviews, checking submitted interviews against designated research sites.

Analysis

Main Analysis

We will study what combination of cash incentives and text/voice reminders affect the intention to get a tuberculosis screening (Intention) and actually take the tuberculosis test (Behavior). We model a binary outcome which is tuberculosis testing (or intention to test) as a function of five treatment conditions (with the Placebo condition as the reference category) using a random effects logistic regression model that accounts for clustering by villages. Our design allows us to do this first by regressing actual tuberculosis testing on the set of treatment condition dummies. The most basic specification is regressing whether the subject is tested for tuberculosis (*Behavior*) and intention to test (*Intention*) on individual treatment status:

$$\text{Behavior}_{ij} = \beta_0 + \beta_{2a}\text{T2a}_{ij} + \beta_{2b}\text{T2b}_{ij} + \beta_{2c}\text{T2c}_{ij} + \beta_{2d}\text{T2d}_{ij} + \omega\mathbf{X}_{ij} + b_j + \epsilon_{ij}. \quad (3)$$

$$\text{Intention}_{ij} = \beta_0 + \beta_{2a}\text{T2a}_{ij} + \beta_{2b}\text{T2b}_{ij} + \beta_{2c}\text{T2c}_{ij} + \beta_{2d}\text{T2d}_{ij} + \omega\mathbf{X}_{ij} + b_j + \epsilon_{ij}. \quad (4)$$

where:

- i and j index individuals and villages respectively
- Behavior_{ij} has a value of 1 if subject i is tested for tuberculosis the week following the video intervention.
- Intention_{ij} has a value of 1 if subject i indicates, after they receive the video treatment, they intend to get tested for tuberculosis at the village testing clinic the following week.
- β_0 is the placebo probability of getting tested for tuberculosis in the week following the intervention
- β_{2a} is the treatment effect of watching the Health message about tuberculosis testing with NO cash incentive and No text/voice reminders.
- β_{2b} is the treatment effect of watching the Health message about tuberculosis testing with NO cash incentive and with text/voice reminders.
- β_{2c} is the treatment effect of watching the Health message about tuberculosis testing with a cash incentive and with NO text/voice reminders.
- β_{2d} is the treatment effect of watching the Health message about tuberculosis testing with a cash incentive and with text/voice reminders.
- $\mathbf{X}_{ij}$ are covariate controls such as age, gender and education at the individual level.
- b_j is the random village intercept.
- ϵ_{ij} is the error term that is i.i.d. with zero mean.

Given the large sample size, and for ease of interpretation, we will model the absolute probability using a linear probability model. We can assess, as a robustness test, whether we obtain similar results when more deliberately modelling the binary outcome using logistic regression.

Spillover Effects

The cluster randomized trial is designed to identify three spillover effects.

Within Household Spillover Within each sampled household we treat a randomly selected household member. In addition, we identify additional household members. We simply include their names on the survey questionnaire. In post-treatment we will be able

to identify whether or not these additional household members received a TB screening. Within-household spillover is measured by comparing screening rates for these additional, non-treated, household members in the four treatment arms with those untreated additional household members in the placebo arm.

$$\text{Behavior}_{ij} = \beta_0' + \beta_{2a'}\text{T2a'}_{ij} + \beta_{2b'}\text{T2b'}_{ij} + \beta_{2c'}\text{T2c'}_{ij} + \beta_{2d'}\text{T2d'}_{ij} + \omega\mathbf{X}_i + b_j + \epsilon_{ij}. \quad (5)$$

where:

- i and j index individuals and villages respectively
- β_0' is the probability that additional household members of placebo treated households get tested for tuberculosis in the week following the intervention
- $\beta_{2a'}$ is the spillover effect for additional household members in households where the treated individuals watch the Health message about tuberculosis testing with NO cash incentive and No text/voice reminders.
- $\beta_{2b'}$ is the spillover effect for households with the Health message about tuberculosis testing with NO cash incentive and with text/voice reminders.
- β_{2c} is is the spillover effect for households with the Health message about tuberculosis testing with a cash incentive and with NO text/voice reminders.
- β_{2d} is the is the spillover effect for households with the Health message about tuberculosis testing with a cash incentive and with text/voice reminders.
- b_j are village-level fixed effects.

Placebo Within Village Spillover The following regression estimates spillover effects by including the four "embedded placebo" dummy variables: T2a-Placebo, T2b-Placebo, T2c-Placebo and T2d-Placebo. We hypothesize that TB screening rates of subjects receiving the Placebo treatment in non-Placebo villages would be higher than Placebo subjects. In this specification we are testing whether the 25% of Placebo subjects in non-Placebo villages have

screening rates that are different than those for subjects in the villages assigned to the Placebo treatment. Our expectation is that the overall average screening rates will be higher for the 25% in the T2a through T2d treated villages – hence the expectation that $\beta_{3a} > 0$; $\beta_{3b} > 0$; $\beta_{3c} > 0$; and $\beta_{3d} > 0$.

$$\begin{aligned}\text{Behavior}_{ij} = \beta_0 + \beta_{2a}\text{T2a}_{ij} + \beta_{3a}\text{T2a-Placebo}_{ij} + \beta_{2b}\text{T2b}_{ij} + \beta_{3b}\text{T2b-Placebo}_{ij} + \\ \beta_{2c}\text{T2c}_{ij} + \beta_{3c}\text{T2c-Placebo}_{ij} + \beta_{2d}\text{T2d}_{ij} + \beta_{3d}\text{T2d-Placebo}_{ij} + \omega\mathbf{X}_{ij} + b_j + \epsilon_{ij}. \quad (6)\end{aligned}$$

Another version of this analysis that we implement, and should give us identical results, restricts the regression analysis to the placebo subjects and estimate the following equation:

$$\begin{aligned}\text{Behavior}_{ij} = \beta_0 + \beta_{3a}\text{T2a-Placebo}_{ij} + \beta_{3b}\text{T2b-Placebo}_{ij} + \\ \beta_{3c}\text{T2c-Placebo}_{ij} + \beta_{3d}\text{T2d-Placebo}_{ij} + \omega\mathbf{X}_{ij} + b_j + \epsilon_{ij}. \quad (7)\end{aligned}$$

Spillover Untreated Effect. For each treated village clusters we will know the village's population size; the total number of no-treated residents that participated in the TB screening. We will estimate the following count model as a Poisson with a log link equation to test our spillover conjectures (using log population size as an offset):

$$\text{Behavior}_j = \beta_{u0} + \beta_{u2a}\text{UT2a}_j + \beta_{u2b}\text{UT2b}_j + \beta_{u2c}\text{UT2c}_j + \beta_{u2d}\text{UT2d}_j + \omega(\text{log of population})_j + \epsilon_j. \quad (8)$$

where:

- j indexes villages
- Behavior_j is the number of residents tested for tuberculosis the week following the video intervention.
- β_{U2a} is the treatment effect for untreated subjects in villages that were treated with the

Health message about tuberculosis testing with NO cash incentive and No text/voice reminders.

- β_{U2b} is the treatment effect for untreated subjects in villages that were treated with the Health message about tuberculosis testing with NO cash incentive and with text/voice reminders.
- β_{U2c} is the treatment effect for untreated subjects in villages that were treated with the Health message about tuberculosis testing with a cash incentive and with NO text/voice reminders.
- β_{U2d} is the treatment effect for untreated subjects in villages that were treated with the Health message about tuberculosis testing with a cash incentive and with text/voice reminders.
- **(population)** is the village population.
- ϵ_j is the error term that is i.i.d. with zero mean.

Again, our pre-registered hypothesis was that $\beta_{U2c} < 0$ and $\beta_{U2d} < 0$ – those untreated subjects most proximate to the cash incentive treatment would have relatively depressed screening rates.

Within Subject Spillover The trial is designed to estimate the within-subject spillover effect of the treatment arms outlined in Table 1. Six months after the initial video treatment concerning TB screening, subjects receive a follow-up video treatment that simply presents the health benefits of hypertension screenings. Hypertension testing clinics are set up in their villages one-week after this intervention. To measure the effect of the initial treatment on subsequent compliance with a health protocol, we estimate the following reduced-form regression:

$$\text{Behavior2}_{ij} = \beta_0 + \beta_{2a}\text{T2a}_{ij} + \beta_{2b}\text{T2b}_{ij} + \beta_{2c}\text{T2c}_{ij} + \beta_{2d}\text{T2d}_{ij} + \omega\mathbf{X}_{ij} + b_j + \epsilon_{ij}. \quad (9)$$

Behavior2 indicates whether or not the subject received a hypertension screening. Our pri-

mary hypothesis is that subjects in the TB cash treatments (T2c and T2d) will have lower hypertension screening rates than those in the original no cash TB treatments (T2a and T2b) and lower than those originally treated with the placebo (T1).

Additional Sources of Individual treatment effect heterogeneity

We will observe TB testing behavior at the individual-level. Therefore we have the opportunity to understand how our messaging interventions vary across demographic characteristics of our subjects and their locale. This analysis could, of course, have implications for enhancing the effectiveness of TB testing campaigns. Our potential sample of approximately 9,000 subjects will include quite diverse demographic profiles; and the villages themselves will vary quite substantially in terms of size, socio-demographics and administrative features.

The covariate measures obtained in the survey are: gender, age, consumption expenditures, self-assessed financial situation, education, household size and number of children, and employment status.

We do not have strong priors as to precisely which covariates, either individual-level socio-demographic features or village-level aspects, will be the source for heterogeneous treatment effects. We will therefore exploit machine learning techniques to identify likely sources of heterogeneous treatment effects. We will use Bayesian Additive Regression Trees (BART), a non-parametric forest-based estimation strategy, to estimate the relationship between non-randomised features of our observations, subjects' treatment status, and TB testing rates (Duch et al., 2020; Green and Kern, 2012; Hill, 2011). In a recent working paper (Robinson et al., 2023) we extended this BART strategy to allow us to identify covariate profiles that best partition individual-level treatment effects. We will adapt this strategy for this study by identifying which groups of individuals (defined by socio-demographic and geographic features) for which the treatment interventions are most (and least) effective. This analysis will allow us to propose optimal targeting strategies for identifiable segments of the Ghana population.

Balance and Attrition

To assess balance on covariates we will compare their standardized mean differences (raw differences in proportion for binary variables) across the four treatment arms. We will also compare these unadjusted differences with those obtained when the sample is weighted using propensity score matching. The detailed results and comparisons will be generated by the R programme, *cobalt*, (Greifer, 2022).

Our goal in Post-treatment Spill-over (Phase 3) is to re-contact as many of the 9,000 subjects who were treated in Phase I. We do anticipate attrition between the two waves. There is a six-month period separating the initial and the follow-up contact with participants. We have a complete set of demographic measures from the Phase I survey. Comparing the demographic profiles of subjects assigned to the different treatment arms will signal whether missingness is, or is not, independent of treatment status. To the extent that we observe any significant attrition in the sample and evidence that it might be correlated with treatment status, we will implement estimation strategies designed to address any bias that might be associated with this attrition (Hogan, Roy and Korkontzelou, 2004).

Inverse probability weighting would be one of the estimation strategies we would implement. It has the value of being very straight-forward – essentially modeling the attrition process as a function of observable covariates (Anderson et al., 2021). The weights are based on the predicted values from a logistic regression of a binary variable indicating whether the observation is missing on the available covariates, where all the available covariates are allowed to interact with the treatment indicators.[2] The weight is simply 1 over 1 minus these predicted probabilities.[3] We would then re-estimate treatment effects on the subset of the data where outcomes are observed and weight that estimate using these weights. The estimate from this regression is a consistent estimate for the treatment effect assuming the censoring process is observable. An alternative estimation strategy that we will consider is the bounding estimator proposed by (Lee, 2009). He assumes that attrition is monotonic, which in our setting, implies

[2] An alternative here, that we would consider using, is a propensity score estimation algorithm to fully model any possible nonlinearities – the *twang* package, for example (Ridgeway et al., 2021).

[3] These weights are often characterized as "unstable" – a slightly modified estimation strategy can generate more "stable" weights.

that any subject who would not be missing if assigned to the control group would also not be missing if assigned to one of the treatment arms.

Note that we also anticipate attrition in the two text/voice messaging treatments – T2b and T2d. The text/voice messages are sent to treatment subjects three days before, and on the day of, the TB testing clinics. There will be about a 10 day lag between treatment and when the text/voice messages are sent. We anticipate minimal attrition during this period (or cell phone numbers that are not working). Nevertheless we will conduct similar attrition tests for these subjects.

Blinding

The data analytics will be conducted employing a "blinding" protocol (MacCoun and Perlmutter, 2015). One investigator will be responsible or "blinding the data" – essentially perturbing the data before teams investigators conduct analyses of the data – they analyze the data without knowing the blinding conditions. Our strategy will be to have the real data interspersed with other data that has been perturbed. Each researcher would be blinded to whether they were analyzing the real data or the perturbed data.

Ethics

The experiment is conducted according to the University of Oxford's policy for human subjects research. The experiment was approved by the University of Oxford Social Sciences Department of Economics Research Ethics Committee (DREC) with reference R82364/RE002. The experiment has also received ethics approval from the Ghana Health Service Ethics Committee (GHS ERC Approval Notification: Protocol ID NO: GHS-ERC 016/07/23). Informed consent is obtained from each participant at the beginning of the survey.

Pilot Study

We will implement a pilot for the full study. The pilot will be a scaled-down version of the design described above. We will test for fewer treatment effects. The reduced sample size will be sufficiently large such that we will be appropriately powered to detect the hypothesized treatment effects.

Table 2 summarizes the design features of the pilot study. There are three treatment arms in the pilot study:

- *TB Health*: A 45 second TB promotional and information video (based on the English NHS videos);
- *TB Health + Text*: A 45 second TB promotional and information video (based on the English NHS videos) followed (three days before the screening) with a reminder voice/text message;
- *TB Health + $3.00*: A 45 second TB promotional and information video (based on the English NHS videos) that includes a 15 second message offering $3.00 if the respondent receives the TB screening scheduled for their village.

Table 2: TB Screening Pilot Study Design

	District 1		District 2		Total	
	Subjects	Villages	Subjects	Villages	Subjects	Villages
TB Health	338	13	338	13	676	26
TB Health + $3.00	338	13	338	13	676	26
TB Health + Text	338	13	338	13	676	26
Total	1014	39	1014	39	2028	78

The pilot will be conducted in two districts in Ghana: Cape Coast Metro and Akwapim South. The implementation of the interventions will strictly follow the protocol described above for the main study:

- The 78 villages will be selected randomly according to the village selection method outlined above – here we will have triplets rather than quintuplets;

- households and subjects will be randomly selected according to the protocol described above;
- interviews and video treatment interventions will take place on tablets as described earlier;
- the voice/text messaging will follow the protocol described above;
- the TB screening clinics will take place in the villages according to the earlier described protocol
- verified TB screenings will be collected as described above.

The primary hypothesis is the following:

- TB Screening rates for those in the *TB Health + $3.00* treatment arm will be higher than rates for subjects in the simple *TB Health* video treatment arm

Our secondary hypotheses are the following:

- TB Screening rates for those in the *TB Health + Text* treatment will be higher than rates for those simple *TB Health* video treatment arm
- TB Screening rates for those in the *TB Health + $3.00* treatment arm will be higher than rates for subjects in the *TB Health + Text* treatment

Our village cluster design with a total of 78 villages and 2028 subjects is powered at 0.80 to detect an effect size of 0.05.

References

Anderson, Kaitlin, Gema Zamarro, Jennifer Steele and Trey Miller. 2021. "Comparing Performance of Methods to Deal With Differential Attrition in Randomized Experimental Evaluations." *Evaluation Review* 45(1-2):70–104.

Arezki, Rabah. 2021. "Cash payments in Africa could boost vaccine uptake: When doses finally arrive, use cash incentives to overcome hesitancy." *Nature* 9(596).

Athey, Susan, Kristen Grabarz, Michael Luca and Nils C. Wernerfelt. 2022. The Effectiveness of Digital Interventions on COVID-19 Attitudes and Beliefs. NBER Working Papers 30273 National Bureau of Economic Research, Inc.

Benjamin-Chung, Jade, Benjamin F Arnold, David Berger, Stephen P Luby, Edward Miguel, John M Colford Jr and Alan E Hubbard. 2018. "Spillover effects in epidemiology: parameters, study designs and methodological considerations." *International Journal of Epidemiology* 47(1):332–347.

Benjamin-Chung, Jade, Jaynal Abedin, David Berger, Ashley Clark, Veronica Jimenez, Eugene Konagaya, Diana Tran, Benjamin F Arnold, Alan E Hubbard, Stephen P Luby, Edward Miguel and Jr Colford, John M. 2017. "Spillover effects on health outcomes in low- and middle-income countries: a systematic review." *International Journal of Epidemiology* 46(4):1251–1276.

Bennett, S., T. Woods, W.M. Liyanage and D.L. Smith. 1991. "A simplified general method for cluster-sample surveys of health in developing countries." *World Health Statistics Quarterly* 44(3).

Bosu, William Kofi and Dary Kojo Bosu. 2021. "Prevalence, awareness and control of hypertension in Ghana: A systematic review and meta-analysis." *PLOS ONE* 16(3):1–37.

Brewer, Noel T., Alison M. Buttenheim, Chelsea V. Clinton, Michelle M. Mello, Regina M. Benjamin, Timothy Callaghan, Arthur Caplan, Richard M. Carpiano, Renee DiResta, Jad A. Elharake, Lisa C. Flowers, Alison P. Galvani, Peter J. Hotez, Rekha Lakshmanan, Yvonne A. Maldonado, Saad B. Omer, Daniel A. Salmon, Jason L. Schwartz, Joshua M. Sharfstein and Douglas J. Opel. 2022. "Incentives for COVID-19 vaccination." *The Lancet Regional Health - Americas* 8:100205.

Dai, Hengchen, Silvia Saccardo, Maria A. Han, Lily Roh, Naveen Raja, Sitaram Vangala, Hardikkumar Modi, Shital Pandya, Michael Sloyan and Daniel M. Croymans. 2021. "Behavioural nudges increase COVID-19 vaccinations." *Nature* 597(7876):404–409.

Duch, Raymond, Denise Laroze, Thomas Robinson and Pablo Beramendi. 2020. "Multi-modes for Detecting Experimental Measurement Error." *Political Analysis* p. 1–21.

Duch, Raymond, Edward Asiedu, Ryota Nakamura, Thomas Rouyard, Alberto Mayol, Adrian Barnett, Laurence Roope, Mara Violato, Dorcas Sowah, Piotr Kotlarz and Philip Clarke. 2023. "Financial incentives for COVID-19 vaccines in a rural low-resource setting: a cluster-randomized trial." *Nature Medicine* 29(12):3193–3202.

Ghana Health Service. 2018. "The Health Sector in Ghana: Facts and Figures 2018.".

Gibson, Dustin G., Benard Ochieng, E Wangeci Kagucia, Joyce Were, Kyla Hayford, Lawrence H Moulton, Orin S Levine, Frank Odhiambo, Katherine L O'Brien and Daniel R Feikin. 2017. "Mobile phone-delivered reminders and incentives to improve childhood immunisation coverage and timeliness in Kenya (M-SIMU): a cluster randomised controlled trial." *The Lancet Global Health* pp. 428–438.

Giles, Emma L., Shannon Robalino, Elaine McColl, Falko F. Sniehotta and Jean Adams. 2014. "The Effectiveness of Financial Incentives for Health Behaviour Change: Systematic Review and Meta-Analysis." *PLOS ONE* 9(3):1–16.

Green, Donald P. and Holger L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76(3):491–511.

Greifer, N. 2022. cobalt: Covariate Balance Tables and Plots. R package version 4.4.1. Technical report.

Hill, Jennifer L. 2011. "Bayesian Nonparametric Modeling for Causal Inference." *Journal of Computational and Graphical Statistics* 20(1):217–240.

Hogan, Joseph W., Jason Roy and Christina Korkontzelou. 2004. "Handling drop-out in longitudinal studies." *Statistics in Medicine* 23(9):1455–1497.

Institute for Health Metrics and Evaluations. 2018. "Global Burden of Disease. Tuberculosis Situation in 2019. Ghana: IHME.".
URL: *https://vizhub.healthdata.org/gbd-compare/*

Kuupiel, Desmond, Benjamin S. N. Cheabu, Peter Yeboah, James Duah, Joseph K. Addae, Ignatius T. Ako-Nnubeng, Francis A. Osei, Shamsu-Deen Ziblim, Gugu G. Mchunu, Julian D. Pillay and Vitalis Bawontuo. 2023. "Geographic availability of and physical accessibility to tuberculosis diagnostic tests in Ghana: a cross-sectional survey." *BMC Health Services Research* 23.

Lagarde, M., A. Haines and N. Palmer. 2009. The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. Technical Report CD008137.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76(3):1071–1102.

MacCoun, Robert and Saul Perlmutter. 2015. "Blind analysis: Hide results to seek the truth." *Nature* 526.

Mardi, Parham, Shirin Djalalinia, Reza Kargar, Mahnaz Jamee, Zahra Esmaeili Abdar and Mostafa Qorbani. 2022. "Impact of incentives on COVID-19 vaccination; A systematic review." *Frontiers in Medicine* 9.

Merriam, Sarah and Hannah Behrendt. 2020. Increasing vaccine uptake in low- and middle-income countries: Opportunities for behavioural insights research. Technical report Behavioural Insights Team.

Milkman, Katherine L., Linnea Gandhi, Mitesh S. Patel, Heather N. Graci, Dena M. Gromet, Hung Ho, Joseph S. Kay, Timothy W. Lee, Jake Rothschild, Jonathan E. Bogard, Ilana Brody, Christopher F. Chabris, Edward Chang, Gretchen B. Chapman, Jennifer E. Dannals, Noah J. Goldstein,

Amir Goren, Hal Hershfield, Alex Hirsch, Jillian Hmurovic, Samantha Horn, Dean S. Karlan, Ariella S. Kristal, Cait Lamberton, Michelle N. Meyer, Allison H. Oakes, Maurice E. Schweitzer, Maheen Shermohammed, Joachim Talloen, Caleb Warren, Ashley Whillans, Kuldeep N. Yadav, Julian J. Zlatev, Ron Berman, Chalanda N. Evans, Rahul Ladhania, Jens Ludwig, Nina Mazar, Sendhil Mullainathan, Christopher K. Snider, Jann Spiess, Eli Tsukayama, Lyle Ungar, Christophe Van den Bulte, Kevin G. Volpp and Angela L. Duckworth. 2022. "A 680,000-person megastudy of nudges to encourage vaccination in pharmacies." *Proceedings of the National Academy of Sciences* 119(6):e2115126119.

Mobarak, Ahmed Mushfiq, Niccolò Meriggi, Maarten Voors, Madison Levine, Vasudha Ramakrishna, Desmond Maada Kangbai, Michael Rozelle, Ella Tyler and Sarah Cundy. 2022. "Solving Last-Mile Delivery Challenges is Critical to Increase COVID-19 Vaccine Uptake: A Cluster Randomized Controlled Trial.".

Moehring, Alex, Avinash Collis, Kiran Garimella, M. Amin Rahimian, Sinan Aral and Dean Eckles. 2021. "Providing normative information increases intentions to accept a COVID-19 vaccine.".

Perneger, Thomas V. 1998. "What's wrong with Bonferroni adjustments." *BMJ* 316(7139):1236–1238.

Reñosa, Mark Donald C, Jeniffer Landicho, Jonas Wachinger, Sarah L Dalglish, Kate Bärnighausen, Till Bärnighausen and Shannon A McMahon. 2021. "Nudging toward vaccination: a systematic review." *BMJ Global Health* 6(9).

Ridgeway, Greg, Dan McCaffrey, Andrew Morral, Matthew Cefalu, Lane Burgette, Joseph Pane and Beth Ann Griffin. 2021. Toolkit for Weighting and Analysis of Nonequivalent Groups: A guide to the twang package. R package RAND.
URL: *https://cran.r-project.org/web/packages/twang/vignettes/twang.pdf*

Robinson, Thomas, Niek Tax, Richard Mudd and Ido Guy. 2023. "Active learning with biased non-response to label requests.".

Sasaki, Shusaku, Tomoya Saito and Fumio Ohtake. 2022. "Nudges for COVID-19 voluntary vaccination: How to explain peer information?" *Social Science Medicine* 292:114561.

Schneider, Florian H., Pol Campos-Mercade, Stephan Meier, Devin Pope, Erik Wengström and Armando N. Meier. 2023. "Financial incentives for vaccination do not have negative unintended consequences." *Nature* 613(7944):526–533.

Sinclair, Betsy, Margaret McConnell and Donald P. Green. 2012. "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments." *American Journal of Political Science* 56(4):1055–1069.

Appendix

Code for power analysis

The following code was used to estimate the power of the experimental design, and was run on a 2021 Macbook Pro M1 Max running R 4.3.1 on MacOS Ventura 13.6.

```
library(tidyverse)
library(foreach)
library(doParallel)

#### Function ####
re_sd <- function(ICC, e_sd) {
  re_var = (ICC*(e_sd^2))/(1-ICC)
  return(sqrt(re_var))
}

gen_data <- function(subjects = 20,
                     villages = 75,
                     districts = 6,
                     bp = 0.5,
                     icc = 0.015,
                     max_n = 9000,
                     tau_start = 0.2) {

  taus <- c("cash_text" = tau_start,
            "cash_notext" = tau_start*0.75,
            "nocash_text" = tau_start*0.5,
            "nocash_notext" = tau_start*0.25)

  treats <- c("placebo",names(taus))

  ds <- paste0("d",1:districts)
  vs <- paste0("v",1:villages)

  v_res <- rnorm(length(vs), 0, sd = re_sd(icc, sqrt(bp*(1-bp))))

  geog_data <- data.frame(d_id = rep(ds, each = villages),
                          v_id = vs,
                          v_re = v_res) %>%
    group_by(d_id) %>%
    mutate(v_treat = sample(rep(treats,villages/length(treats))))

  subj_data <- data.frame(
    s_id = paste0("s",1:subjects*villages*districts),
    v_id = rep(vs, each = subjects),
    d_id = rep(ds, each = subjects*villages)
  ) %>%
    left_join(geog_data, by = c("d_id","v_id")) %>%
    group_by(d_id, v_id) %>%
    mutate(s_treat = sample(c(rep("placebo",0.25*subjects), rep("noplacebo", 0.75*subjects)))) %>%
    mutate(treat = ifelse(v_treat != "placebo" & s_treat == "placebo", "placebo", v_treat))

  subj_data$p <- bp +
    taus["nocash_notext"]*I(subj_data$treat == "nocash_notext") +
    taus["nocash_text"]*I(subj_data$treat == "nocash_text") +
    taus["cash_notext"]*I(subj_data$treat == "cash_notext") +
    taus["cash_text"]*I(subj_data$treat == "cash_text") +
    subj_data$v_re

  subj_data$p <- case_when(subj_data$p > 1 ~ 1,
                           subj_data$p < 0 ~ 0,
                           TRUE ~ subj_data$p)

  subj_data$y <- rbinom(nrow(subj_data),1,subj_data$p)
```

```
  return(subj_data)

}


simulate <- function(bp,
                     tau_start = 0.2,
                     lambda = 0.95,
                     power = 0.8,
                     sims = 500,
                     coef = NA) {

  if (lambda > 1 | lambda < 0) {
    stop("lambda must be between 0 and 1")
  }

  if (is.na(coef)) {
    stop("must provide regression coefficient to assess")
  }

  n.cores <- parallel::detectCores() - 1

  my.cluster <- parallel::makeCluster(
    n.cores,
    type = "PSOCK"
  )
  doParallel::registerDoParallel(cl = my.cluster)
  foreach::getDoParRegistered()

  powered = TRUE
  min_beta = tau_start
  tau_ct = tau_start

  while (powered) {

    cat(paste0("\r Testing tau of ",tau_ct))

    results <- foreach(i = 1:sims,
                       .combine = 'c',
                       .packages = 'tidyverse',
                       .export = c("gen_data","re_sd")) %dopar% {

      sim_data <- gen_data(bp = bp,
                           tau_start = tau_ct,
                           max_n = max_n)

      sim_data$treat <- relevel(as.factor(sim_data$treat), ref = "placebo")

      res <- summary(
        lm(y ~ treat + as.factor(v_id),
           data = sim_data)
      )

      res$coefficients[coef,"Pr(>|t|)"] < 0.05
    }

    powered <- mean(results) > power

    if (powered) {
      min_beta = tau_ct
      tau_ct <- tau_ct*lambda # scale the hypothesised effects
    }

  }

  parallel::stopCluster(cl = my.cluster)
```

```
  return(min_beta)
}

#### Main analysis ####
bps <- seq(0.5,0.9,0.05)
min_effects <- data.frame(bp = bps,
                          es = NA)
set.seed(89)
for (bp in bps) {
    min_effects$es[min_effects$bp == bp] <- simulate(
        bp = bp,
        tau_start = 0.24,
        coef = "treatnocash_notext",
        sims = 1000
    )
}

# divide by 4 as looking at min treat group
ggplot(min_effects, aes(x = bp, y = es*0.25)) +
  geom_line(size = 1, color = "firebrick1") +
  geom_point(color = "grey15", size = 2) +
  labs(x = "Baseline probability of getting tested",
       y = "Minimum effect size (for each treatment arm)") +
  theme_minimal() +
  ylim(0,max(min_effects$es)*0.25 + 0.01) +
  theme(text = element_text(size = 14))
```