Analysis Plan

Project Name: Increasing Effectiveness of Family Planning Promoters in Mozambique through an SMS Intervention

Project Code: 1811-A

Date Updated: February 3, 2020

Project Description

OES is supporting the USAID-funded Integrated Family Planning Program (IFPP) in Mozambique led by Pathfinder International. The component of the IFPP that will be evaluated by OES is led by Population Services International (PSI). It aims to reach and engage urban women in Sofala and Nampula provinces at their doorstep with tailored family planning counseling specific to their needs, provided by family planning promoters. OES is supporting USAID/Mozambique efforts to increase the effectiveness of IFPP family planning promoters by providing them with additional tools to facilitate counseling beneficiaries and providing clinic referrals. USAID/Mozambique is interested in learning more about different strategies to make the IFPP promoters as effective as possible, including but not limited to supplemental messaging directed at beneficiaries or promoter job aides. The aim is to increase the impact of the Agency's development dollars. OES will also support USAID/Mozambique to measure the impact of these additional tools, through a rigorous evaluation.

The goal of the study is to use a randomized design to evaluate the effect of sending a series of mobile SMS follow-ups to beneficiaries who have received referrals for family planning services from an IFPP family planning promoter. The activity seeks to understand whether behaviorally informed SMS messages can be a cost-effective means of increasing the referral redemption rate among beneficiaries.

Data and Data Structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to the data structure and variables.

Primary data is collected at the individual beneficiary level via the mobile Connecting with Sara (CwS) application, which PSI-Mozambique currently uses to track clinic referrals and redemptions. Data on referrals and basic demographics are collected at an individual level by the promoter; data on redemptions are collected at an individual level by program-trained family planning nurses based at public health clinics. The nurses enter redemption codes into the application to indicate

that a woman has come in for a consultation and records the result of the visit (i.e., method issued). These data are stored in the District Health Information System (DHIS2), an open source health management information system and data warehouse, which is updated daily. This dataset includes at minimum the following variables:

- *Marieta ID* -- beneficiary identification code based on location and birth order, but not otherwise linkable to personal identity;
- *Date* -- date and time of transaction recorded, used to calculate the time between receipt and redemption of a referral;
- *Longitude and Latitude* -- coordinates as recorded at the time of data upload by the GIS functionality of the mobile phone being utilized;
- *Provider - New or Continuing Users* -- categorical variable classifying beneficiaries who redeem referrals as continuing users, method switchers, or new users;
- *Phone Ownership* -- categorical variable recording who owns the phone as described by the beneficiary when receiving a referral from a promoter (i.e.,owned or shared);
- *Actor Type* -- categorical variable denoting the role of the individual uploading the information (i.e.,promoter or nurse);
- *Marieta - Age* -- categorical variable indicating the age range of the beneficiary in five-year increments between under 15 to 50 and over;
- *Geography - Province* -- province identifier indicating Sofala or Nampula;
- *Transaction Type - Acronym* -- categorical variable denoting the session type (i.e.,promoter session with referral, promoter session without referral, reminder voucher issued by the promoter, or nurse-recorded redemption);
- *CwS - FPL - Service Received* -- categorical variable denoting the type of method issued by the nurse (i.e.,pill, implant, injection, or counseling);
- *Program* -- program identifier for the recorded interaction (i.e.,IFPP, Nurse, Tem+, etc.).

In addition, PSI-Mozambique maintains consent records (i.e., copies of signed consent forms and affiliated phone numbers), output from running the randomization process (i.e., which phone numbers are assigned to the intervention versus control groups), and records of messages sent to the intervention group. The records of messages sent include the time and date of each message sent to each phone number as well as status (i.e., sent or delivered, but not whether the message was opened). These data are stored separately via OneDrive and will be used to construct:

- An indicator denoting which beneficiaries provided consent forms and hence were included in the study sample for randomization,
- An indicator denoting treatment allocation to either the intervention or control group, and
- Variables containing the dates and times of messages issued to the phone numbers assigned to the intervention group, and whether these messages were received.

Since randomization is ongoing, we will monitor and perform checks throughout the data collection process to ensure efficiency and efficacy of intervention protocols (i.e., accuracy of

block randomization at the supervisor level, relay of eligible phone numbers to the SMS vendor, etc.). These checks will include periodic download of data from DHIS2 and individual inspection of consent records, randomization output, and message records.

Data from the DHIS2 will be downloaded for the entire period coinciding with the fieldwork portion of the study and merged with the consent records, randomization output, and message records based on the phone number affiliated with each record to construct the final dataset for analysis purposes. The final dataset used for analysis therefore will be at the individual beneficiary level and will include (1) which beneficiaries were enrolled and assigned to the intervention or control groups respectively, (2) records of referral redemptions for those beneficiaries through the end of the study period, and others as consistent with the analyses below.

Outcome Variables to Be Analyzed:

The primary outcome of interest is the referral redemption rate. The raw data from DHIS2 contains individual records of issued referrals and which of these have been redeemed with nurses to be able to calculate this overall redemption rate. These will be compared between the intervention and control groups as described in the Statistical Models & Hypothesis Tests section. Additional outcomes of interest include time to redemption, which can be calculated by differencing the date/timestamps, and types of methods received.

Transformations of Variables:

Redemptions are recorded at an individual beneficiary level by nurses -- if a beneficiary redeems a referral with a nurse within the study period, she will be coded as redeemed (i.e., "1"). If a beneficiary has no redemption with a nurse recorded, she will be coded as no redemption (i.e., "0"). Time to redemption will be calculated for each referral redeemed by differencing the nurse redemption date/timestamp and the issue date/timestamp contained within the "Date" variable in DHIS2.

While the study design is such that the first SMS message should be within 4 days of referral receipt, there will be some variation in timing across beneficiaries assigned to the intervention group. Precise message timing will be calculated for each beneficiary assigned to the intervention group by differencing the date stamp of the first message sent by the vendor from the message records and the referral issue date stamp contained within the "Date" variable in DHIS2. This calculated variable then can be used to evaluate the presence of differential effects by timing of messages sent relative to referral receipt and possibly determine optimal message timing.

Finally, distance to the clinic will be calculated from the GPS-location of the referral issued by the promoter and the GPS coordinates of the nearest clinic. This can also be calculated by calculating the distance between the GPS location of the referral issued and the GPS location of the nurse recorded redemption respectively for beneficiaries who redeem referrals for comparison.

Imported Variables:

PSI will provide the consent records and randomization output, which are used to construct the indicator denoting assignment to either the intervention or control group, in addition to the timing/status of messages sent as provided by the SMS vendor. These records are maintained at an individual level and will be merged with the DHIS2 dataset based on the phone numbers provided.

Data Exclusion:

Information from individuals who do not provide consent to be part of the study or are not issued a referral are still recorded within the CwS application per standard data collection practices conducted by PSI-Mozambique. However, data access and analysis will pertain only to the sample that provided consent. All enrolled participants who provided consent will be included in the intent-to-treat analysis; any participant for whom we do not have a consent form on file will be excluded from analysis.

Treatment of Missing Data:

Since PSI maintains a comprehensive administrative dataset and our analysis is restricted to the consenting participants, we do not expect true missing values. Any missing values for referral redemptions will be interpreted as absence of redemption (i.e.,"not redeemed"). We will not impute any covariates for those not reporting in the DHIS2 data.

However, data on message delivery will be provided by the SMS vendor that is managing the intervention and must be compiled and maintained by PSI. We are aware there is a significant risk of missing data here. If data is missing, we will code any recipients for whom delivery was not recorded as 1 for message delivery, thus ensuring a conservative estimate of the intervention's implementation.

Descriptive Statistics, Tables, & Graphs

We will plan to present descriptive statistics that may be useful for PSI's implementation, for example: average time to redemption by month, by promoter supervisor, and by location. We do not anticipate at this point that these materials will also be included in the project abstract or other analysis products.

Statistical Models & Hypothesis Tests

This section describes the statistical models and hypothesis tests that will make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models:

Randomization Test
Before continuing with analysis, we will check the initial randomization by conducting balance tests using observable characteristics (i.e., age, clinic proximity) on the full dataset. These balance tests will be reported in the final analysis.

Treatment Effects
We will estimate the intent-to-treat (ITT) effect between the control group (no SMS messages) and the intervention group (SMS messages). We will estimate an ordinary least squares (OLS) regression of the binary outcome (redeemed referral) on intervention assignment for beneficiary *i* in strata *s*. The strata corresponds to a supervisor-date cell: i.e., all women who had been provided with a referral by promoters working under a certain supervisor, who were randomized as of a certain date. We estimate that the full sample will include at least 1,000 strata.[1]

$$y_{is} = \alpha + \beta_1 * SMS_{is} + \beta_1 * Age_{is} + \beta_3 * Provi_s + \beta_4 * Dist_{is} + \varepsilon_{is}$$

In the primary OLS specification, we will regress the referral redemption indicator on:

- An indicator variable denoting assignment to the intervention group (i.e., receipt of SMS messages), and
- Beneficiary and supervisory-group level covariates as available in the data, such as age group, province, and distance to the clinic.

We can repeat this process for other outcomes of interest, such as time to redemption and type of method received. Covariates are included to enhance the precision of our estimate of treatment effect by adjusting for pre-existing differences in characteristics. Standard errors will be adjusted using clustering by randomization day.

In addition, we will estimate an instrumental variables model parallel to the above specification regressing the outcome variable on a dummy variable for $SMSdelivered_{is}$, instrumented by instrumented by SMS_{is}. This will allow us to adjust for variation in the probability that the SMS messages are delivered; some SMS may not be delivered if there are network challenges, the phone is switched off, or the number reported is invalid or becomes invalid. *SMSdelivered* will be coded as zero if the first two messages are not reported delivered, and one if the first two messages are reported delivered; this reflects our ex-ante hypothesis that receipt of the first two messages is most important.

In addition to the OLS specification, we will run an analogous logistic regression given that much of the related literature, particularly in health journals, reports logistic results. If the OLS and logistic regressions yield substantively different results, the OLS results will take precedence.

[1] A strata is defined as a supervisor group X randomization day. We estimate that the evaluation period will include around 50 randomization days (twice-weekly randomization over a period of approximately six months), and the sample includes 20 supervisors.

In addition, we plan to estimate the primary specification for a restricted sample excluding the first month of enrollees (from February 1 to March 1). We anticipate the study will have a "soft launch" while randomization procedures become more regularized in the first month. The soft launch period may be characterized by a longer and more variable gap between enrollment and first message for beneficiaries enrolled in the first month. Given that randomization is independent on each day on which randomization is conducted, the sample can be restricted while still preserving the randomization within the subsample that enters the evaluation after March.

Follow-Up Analyses:

Subgroup analysis

We will conduct several types of subgroup analysis.

1) *Timing effects*

While the study design is such that the first SMS message should be within 4 days of referral receipt, there will be some variation in timing across beneficiaries assigned to the intervention group. We therefore also can examine the intervention effects on referral redemption for the subgroup that received the message more quickly (i.e., within 2 days, 4 days, 6 days, etc.).

However, it should be noted that the time elapsed between referral receipt and the first message may be endogenous, and the randomization procedure does not stratify on this variable. Accordingly, this analysis will be considered correlational.

2) *Distance effects*

In addition, we will analyze heterogeneous effects for women with respect to distance from the nearest clinic. This analysis may inform us if the SMS reminders are effective in stimulating take-up of referrals for women who are relatively closer to the clinic.

3) *Beneficiary characteristics*

We will also analyze heterogeneity with respect to beneficiary characteristics (new versus continuing user, and beneficiary age).

4) *Cost-Benefit Analysis*

If SMS reminders are found to have a significantly positive effect on redemption, We will present a tabulation of associated costs of SMS reminders to compare against potential benefits anticipated from expanding the intervention. The analysis will include discussion of available infrastructure to support such operations in a broader context, contextual adaptability (i.e., translation into other languages or for other populations), and assessment of costs from varying perspectives.

Inference Criteria, Including Any Adjustments for Multiple Comparisons:

We will use standard inference criteria without any adjustment for multiple comparisons. We will use two-tailed tests and three threshold significance levels of 0.01, 0.05, and 0.10. We will report estimates as statistically significant if they have a *p*-value less than or equal to 0.05.

Limitations:

One immediate limitation is that the anticipated sample size remains uncertain given challenges in appropriately verifying consent. In order to address this limitation, we have estimated statistical power for three alternate scenarios, in which we receive consent from 90%, 75%, and 60% of eligible women, respectively. In the first scenario, in the project period of seven months we expect to enroll 5100 women; in the second scenario, we would enroll 4700; in the third scenario, we would enroll 3700. The minimum detectable effect in percentage points then varies from 3.9 to 4.6 percentage points. We also recognize that we may observe a final sample size that is outside these estimated parameters.

Another limitation is that we may have identified the wrong barrier limiting redemption of referrals. This evaluation was designed based on the hypothesis that low redemption rates of referrals reflect the diminishing salience of promoter messaging over time; if beneficiaries do not redeem their referrals quickly, beneficiaries either forget or lose the referral and do not redeem it later. Therefore, it is assumed that the SMS serves a useful reminder to highlight the pending referral, encouraging beneficiaries to go to a clinic if they have not already done so. However, there is the possibility of a null result if low take-up reflects other barriers facing beneficiaries. We discuss some of these alternatives below.

- One plausible alternative is that demand for family planning may not be low in actuality; instead, it may be a byproduct of social desirability bias. Beneficiaries may be agreeing to receive a referral in response to social norms or desire to please the promoter following their interaction rather than having an actual demand for a family planning method or more formal counseling. However, based on available data, about half of all first-visits with promoters do not result in referrals being issued; in November 2019, only about 33% of visits resulted in a referral being issued, and in the pilot, about 60% of recorded interactions declined referrals. These proportions indicate that providing a negative response to receiving a referral is socially acceptable, or at least not socially abhorrent, making this alternative hypothesis unlikely.
- Beneficiaries also may not be able to read the message content for it to function as a reminder, especially since the literacy rate for the country is fairly low (around 60%). Nevertheless, the preliminary focus group testing and piloting conducted thus far suggests that eligible beneficiaries generally are able both to read and understand the message content. Coupled with the fact that beneficiaries provide consent to receive SMS messages, it is reasonable to assume that the vast majority will be able to read and understand the short messages.

- Finally, beneficiaries may be interested in family planning services, but the costs (i.e., time spent traveling and attending the nurse visit at the clinic) are too high such that the reminder message is insufficient. One way in which we can assess this alternative is via the analysis of heterogeneous effects by distance. It may be that the reminder is sufficient only up to a particular cost/distance threshold and that the efficacy diminishes beyond that radius.

Collectively, these alternatives may be explored further depending upon preliminary results from the trial. For example, there is potential for an additional follow-up phone survey since beneficiaries have consented to contact by phone for this study. This type of data collection has been utilized by PSI-Mozambique in the past, and so there is precedent for its management and coordination. We can supplement data collection on other potential barriers like social norms, receipt and comprehension of SMS messages, and related travel costs in this way. Doing so would also provide additional user feedback about the intervention. Most likely, this would be conducted in July with a subsample of the enrolled participants or cohorts contingent on necessity, availability of supplementary funding, and corresponding IRB amendment approval.

Exploratory Analysis:

We may also conduct exploratory analysis of possible effects on type of method received — even if statistical power is low for these individual methods, this may be done using a Bayesian approach.