PAP Deviation

Christopher Grady, Rebecca Wolfe, Dawop Saidu, Lisa Inks

2023-03-01

Overall

This document describes deviations from the pre-analysis plan listed here.

Change 1: Hypotheses

Description & Justification

We initially listed 7 hypotheses. We now list 4. These changes apply to the community-level analysis and the individual-level analysis.

1. We combined hypotheses 1 ("increase trust and positive affect") and 4 ("increase positive attitudes") for two reasons. First, we dropped survey measures of positive affect. Second, the survey includes multiple attitudinal measures that focus on underlying attitudes regarding trustworthiness and we felt the general "positive attitudes" wording was a more accurate description than the specific "trust" wording.
2. We combined hypotheses 2 ("more likely to cooperate in the PGG") and 7 ("fewer community members donate nothing during the Public Goods Game") because both are about donations in the public goods game.
3. We dropped hypothesis 6 ("Communities that receive the ECPN program will resolve disputes more peacefully compared to control communities."). Hypothesis 6 is beyond the scope of this paper. This paper is about individual attitudes, perceptions, and behaviors. Hypothesis 6 is about the success of dispute resolution. In addition, the dispute resolution aspect of the intervention was targeted at community leaders rather than the whole community. We are considering a separate paper about dispute resolution.

In the individual-level hypotheses, we unintentionally left out the individual-level hypothesis for *perceptions of physical security*. We included that individual-level hypothesis.

Sections Amended

Theory and Hypotheses.

Timeline of change

We made the decision to modify these hypotheses after data analysis.

Change 2: Causal Mechanisms

Description & Justification

We initially included a subsection called "Mechanisms Through Mediating Variables". We have removed that section. We do not investigate causal mechanisms in this paper. We also remove the references to causal mechanisms in the survey data. We did not list any hypotheses about causal mechanisms, so our hypotheses are unchanged.

That section may be used as part of the pre-analyis plan for a paper about the causal mechanisms of intergroup contact.

Sections Amended

Theory and Hypotheses; Survey Data; Survey Question Appendix.

Timeline of change

We made the decision to exclude a section about causal mechanisms after data analysis.

Change 3: Placebo outcomes

Category

Hypothesis testing.

Description and Justification

We initially listed only radio listening as a placebo outcome. Along with radio listening, we also use attitudes towards violence and attitudes towards people from othr religions as placebo outcomes. Attitudes about violence are a good candidate for a "placebo outcome" because intergroup contact should not affect general attitudes about violence, but respondents may feel social pressure to answer violence questions in a desirable way. Attitudes towards people from other religions should not be affected by ECPN but would signal either social desirability bias or a general increase in tolerance towards outgroups. And radio listening should also not be affected by ECPN but is less relevant for how self-reported attitudes change from baseline to endline.

We did not initially list attitudes towards violence as a placebo outcome because the program team was interested to know if the program affected these attitudes. However, we explicitly did not include attitudes towards violence in any hypothesis because we did not expect the program to affect such attitudes and wanted to use them as placebos in our analysis.[1]

[1] In our original PAP submission, we wrote: "... while certain outcomes may be of interest to the program team, the impact evaluation team does not include hypotheses about ECPN changing perceptions of the acceptability of some types of violence, about inter-religious trust, or perceptions of the actors involved in dispute resolution." We also did not expect ECPN to affect trust in members of other religions, but we did not believe it was appropriate as a "placebo" outcome because intergroup contact could generalize to a wider religious outgroup and religion separates farmers and pastoralists in one state. And we did not expect ECPN to systematically affect perceptions of the actors involved in dispute resolution but do not use these perceptions as a placebo outcome because ECPN worked with some actors involved in dispute resolution (the particular actor involved varied by community).

Sections Amended

Overall Study Design of Final Evaluation; Research Design and Data Sources.

Timeline of change

We made the decision to include "attitudes towards violence" as a placebo outcome before endline data collection.

Change 4: Individual-level analysis

Category

Estimation.

Description and Justification

We planned to compare individuals who had been assigned to participate and complied with individuals who were assigned to participate but did not. Instead we compare individuals who *did* participate with individuals who *did not* participate. We made this change due to the possibility that individual-level treatment assignment was not followed.

Sections Amended

Overall Study Design of Final Evaluation.

Timeline of change

We made the decision to compare actual participants with actual nonparticipants during endline data collection.

Change 5: Modify regression equation

Category

Estimation.

Description and Justification

We compare treated communities to control communities using a diffrence-in-differences framework. At the time of writing the PAP, we only knew of the "change scores" strategy where the outcome of interest is the change from baseline to endline. After writing the PAP, we learned the "controlling-for" method where the outcome of interest it the endline outcome and the equation controls for the baseline outcome. The "controlling-for" method is more powerful than the "change scores" method and is unbiased when baseline outcomes are balanced. We therefore use the "controlling-for" method when baseline outcomes are balanced and the "change scores" method when baseline outcomes are not balanced.[2]

[2]We learned about this method thanks to DeclareDesign Blog: https://declaredesign.org/blog/use-change-scores-or-control-for-pre-treatment-outcomes-depends-on-the-true-data-generating-process.html

Sections Amended

Research Design and Data Sources

Timeline of change

We made the decision to use the "controlling-for" method when baseline outcomes are balanced during endline data collection.

Change 6: Behavioral observation data

Category

Estimation/outcomes.

Description and Justification

We collected more behavioral observation data than we anticipated. Instead of collecting behavioral observation data only in treatment sites, we also collected it in control sites. This allows us to compare behavioral changes in treatment sites to behavioral changes in control sites. We added a new estimation procedure for the behavioral observation data. These observations were also not as regular as we intended, which is now noted in the PAP.

Our original PAP planned to observe project committee meetings. We were not able to collect observation data at those meetings. This revised PAP removes references to obsrvational monitoring of the meetings.

Sections Amended

Research Design and Data Sources

Timeline of change

We made the decision to gather observational data in control sites after baseline survey data collection.

Change 7: Survey data outcomes

Category

Outcomes/measurement.

Description and Justification

We initially listed survey data outcomes that were unrelated to our hypotheses (for example, survey questions about causal mechanisms, social cohesion/social norms, conflict history, and resource sharing). We said we would combine many of these questions into indices and analyze the effect of ECPN on those outcomes. We are no longer analyzing outcomes not directly related to our hypotheses; we now only list survey outcomes

that relate to our hypotheses. The main change is that the attitudinal index now only contains questions directly related to attitudes towards the outgroup.

In our original PAP, after listing the survey outcomes, we noted that "Some of these concepts are exploratory for the ECPN impact evaluation" and that we did not have hypotheses for those outcomes. We have now removed that because we no longer list the exploratory outcomes in the PAP.

Our original PAP included a list experiment, but the list experiment failed – the control list had higher scores than the treatment list in most communities, indicating that the difference between the treatment list and control list was not the percentage of people who agreed with the treatment item. We therefore exclude the list experiment from our revised PAP's outcomes. We instead footnote that we conducted a failed list experiment.

Our original PAP described a "randomization experimnt". We have renamed that a "percent experiment" to avoid confusion with the randomized response method of maintaining confidentiality. In addition, as our main outcome, we described using the relationship between more outgroup members (5%-75%) and willingness to join the group. Our prediction was that the percentage of outgroup members has less of an effect after the intervention. However, that analysis assumes that 5% functions as a baseline for how likely an individual is to join a group with a very small % of outgroup members. We do not expect that baseline to change. In our paper, the 5% condition had a low "yes" rate because respondents didn't want to join a group with *any* outgroup members. Joining in the 5% condition was correlated with positive outgroup attitudes. The assumption that the 5% condition is an innocuous baseline was violated. We thus assessed whether the intervention increased respondents' "yes" rate at any % of the outgroup, rather than the difference in "yes" rate as the % of the outgroup increased.

Sections Amended

Research Design and Data Sources

Timeline of change

We made the decision to only analyze outcomes directly related to our hypotheses after data analysis (these outcomes were included in a report to the NGO that implemented ECPN).

We made the decision to exclude the list experiment after data analysis.

We made the decision to analyze the "yes rate" of the percent experiment after data analysis.

Change 8: Removed unorthodox measurement and estimation strategies

Category

Measurement and estimation.

Description and Justification

We were going to try controlling for individual-level covariates in community-level analysis. We planned to do so by (1) predicting individual-level outcomes with individual-level covariates and then (2) aggregating the residuals to create the community-level estimats. This method would have been highly unorthodox and we decided against doing so on the advice of colleagues. It has been removed from the PAP; we may produce a methodological paper testing that method of aggregation.

Our original PAP was also planning to use simultaneous hypothesis testing and ordered hypothesis testing. However, our understanding of these testing procedurs was incorrect. We will not be using them and have removed them from the PAP. However, the expectation that led us to propose ordered hypothesis testing was correct: "we expect the coefficient describing the participant-control comparison to be larger than the coefficient describing the nonparticipant-control comparison, and we expect the coefficient describing the nonparticipant-control comparison to be larger than zero."

Our original PAP also planned to combine our analyses from different levels of analysis (community-level comparisons and individual-level comparisons): "To leverage the p-values we have from multiple analyses, we will first record and combine the p-values from each of the three analyses described above (community-level, participants, and nonparticipants). We will then simulate the possible results we could have observed in each of those studies if there was no effect. To simulate these studies with no effect, we break the relationship between treatment and outcomes by shuffling the treatment indicator (for the community-level study) or the outcome scores (for the individual-level studies). We then have the distribution of p-values we would see if ECPN had no effect in each study. Our outcome of interest is the proportion of this "no effect distribution" that is smaller than our observed group of p-values." We may produce a methodological paper testing that method of inference, but we no longer do it in this paper.

Sections Amended

Analytical Strategy: Combining Outcomes and Ordering Hypothesis Tests

Timeline of change

We made the following decisions during data analysis: (1) to not attempt to control for individual-level characteristics, (2) to not use ordered hypothesis testing, and (3) to not combine p-values across levels of analysis.

Change 9: Survey question list

Category

Measurement.

Description and Justification

Our original PAP's appendix listed every question in the survey. We now only list questions we use in the paper.

Sections Amended

Appendix

Timeline of change

We made the decision to only include survey outcomes directly related to our hypotheses after data analysis (these outcomes were included in a report to the NGO that implemented ECPN).

Change 10: No synthetic control.

Category

Estimation.

Description and Justification

Our original PAP mentioned the possibility of using a synthetic control group as the comparison group for our individual-level comparisons. We decided not to use a synthetic control.

Sections Amended

Estimation – Individual Level Panel Data

Timeline of change

We made the decision to not use a synthetic control before endline data collection.