A Placebo Design to Detect Spillovers from an Education-Entertainment Experiment in Uganda*

Anna Wilke[†], Donald P. Green[‡], Jasper Cooper[§]

January 31, 2019

Abstract

Education-entertainment refers to dramatizations designed to convey information and change attitudes. Buoyed by observational studies suggesting that education-entertainment strongly influences beliefs, attitudes, and behaviors, scholars have recently assessed education-entertainment using rigorous experimental designs in field settings. Studies conducted in developing countries have repeatedly shown the effectiveness of radio and film dramatizations on outcomes ranging from health to voting. One important gap in the literature is estimation of social spillover effects from those exposed to the dramatizations to others in the audience members' social network. In theory, the social diffusion of media effects could greatly amplify their policy impact. The current study uses a novel placebo-controlled design that gauges both the direct effects of the treatment on audience members as well as the indirect effects of the treatment on others in their family and in the community. We implement this design in two large cluster-randomized experiments set in rural Uganda using video dramatizations on the topics of violence against women, teacher absenteeism, and abortion stigma. We find several instances of sizable and highly significant direct effects on the attitudes of audience members, but we find little evidence that these effects diffused to others in the villages where the videos were aired.

*We are grateful to Paul Falzone and Gosia Lukomska from Peripheral Vision International (PVI), who produced the video vignettes, and to Innovations for Poverty Action (IPA), Uganda, which oversaw implementation of the campaign and surveys. We also wish to express our deep sense of gratitude to Cristina Clerici, the project manager, and to Jackline Namubiru and Anthony Kamwesigye, the field managers. Sincere thanks go to Susanne Baltes for her contribution to the design of survey instruments, media intervention, and PAPs, and to Winston Lin for his help designing the randomization scheme and for comments on the PAP. Special thanks to Robert Fleischmann for help with the implementation of the sample selection algorithm. This project received IRB approval from Columbia University (protocol AAAP6500), the Mildmay Uganda Research Ethics Committee (MUREC), and the Uganda National Council for Science and Technology (UNCST). The pre-analysis plans for the midline and endline phases of this study may be found at: http://egap.org/registration/2207 and http://egap.org/registration/2580.

[†]amw2229@columbia.edu PhD. Candidate, Columbia University.

[‡]dpg2110@columbia.edu Professor, Columbia University.

[§]jjc2247@columbia.edu PhD. Candidate, Columbia University.

1 Introduction

Philanthropic groups and human rights organizations routinely deploy media interventions in developing countries to promote pro-social behaviors (Blair, Littman, and Paluck, 2017), increase awareness of beneficial technologies (Heong et al., 2008; Banerjee, Barnhardt, and Duflo, 2017), correct misconceptions that contribute to the spread of disease (Abramsky et al., 2014), or discourage harmful or discriminatory behaviors (Abramsky et al., 2014; Babalola et al., 2006; UNICEF, 2005; UNFPA-UNICEF, 2014). The question is whether, and under what conditions, media campaigns on these topics change beliefs, attitudes, and behaviors.

Although the scholarly literatures on propaganda, public service announcements, and education-entertainment programs trace their origins to the 1930s, only recently have studies rigorously assessed radio or video campaigns deployed in developing countries. Inspired by early observational studies that found media dramatizations to have large effects on audience behavior (Singhal, Rogers, and Brown, 1993; Heatherton and Sargent, 2009), the past decade has seen rapid growth in randomized controlled trials evaluating media campaigns. Paluck (2009) and Paluck and Green (2009) evaluated the effects of an ethnic reconciliation radio soap opera in Rwanda by randomly assigning villages to receive recordings of the soap opera or another program on HIV prevention over the course of one year. The ethnic reconciliation program seemed to have little effect on inter-group attitudes, but its messages did affect perceived norms about inter-ethnic cooperation and listeners' proclivity to take action themselves rather than deferring to authorities. Banerjee, Barnhardt, and Duflo (2017) found that education-entertainment movies dramatizing the benefits of iron-fortified salt in Indian villages where shopkeepers were incentivized to distribute it led to an increase in product usage. Green and Vasudevan (2015) aired short radio vignettes depicting the negative effects of vote-buying multiple times each day immediately before the 2014 Indian national elections and found a decrease in votes cast for reputed vote-buying parties. By contrast, media messages not conveyed through dramatization tend to produce minimal effects. Randomly assigning radio transmitters in the Democratic Republic of Congo to air talk shows on intergroup conflict and cooperation, Paluck (2010) found these shows to have an unexpected corrosive effect on intergroup attitudes, making listeners more mindful of grievances. Galiani, Gertler, and Orsola-Vidal (2012) evaluated the effects of 30 to 50 second encouragements to wash hands aired on randomly selected Peruvian radio stations multiple times each day for approximately one year but found no evidence of an effect on views about hand-washing or hand-washing behaviors. Sixty second health-promotion radio spots aired 6-12 times per day for months in Burkina Faso produced weak effects (Sarrassat et al., 2015). From this small assortment of studies, it appears that instructional messages and talk shows have limited effects but that dramatization may produce changes in certain attitudes and behaviors. This pattern is consistent with observational studies that have traced the consequences of the introduction of mass media and entertainment programs to regions in India (Jensen and Oster, 2009) and Brazil (La Ferrara, Chong, and Duryea, 2012).

Why might education-entertainment be an especially effective way of changing how people think and act? One hypothesis is that dramatization leads the audience to identify with the characters in the story and adopt the perspective of someone from a different background or context (Petty and Cacioppo, 1986; Bandura, 2004). Compared to an audience receiving an explicit argument or direct instructions, an audience engrossed by a dramatization of a social problem is less likely to resist the underlying message, especially when it is expressed by protagonists or conveyed by their behavior.

The influence of education-entertainment in developing countries hinges, of course, on exposure. In a country such as Uganda, where one-third of the population lives on an approximately $2 per day (World Bank Development Indicators 2016), the dearth of televisions and smartphones presents a formidable barrier to exposure to video content.[1] In our sample, only 26% of respondents own a TV and only 20% own a mobile phone with internet connectivity. Although radio is more accessible to the very poor, radio ownership (or radio access via cell phones) is far from universal.

Given the challenges of exposing large segments of the population to education-entertainment, the question is whether such media campaigns have second-hand or spillover effects. If audience members absorb the media message and convey it to family members or others in their social network, the media campaign's reach expands, perhaps by a sizable factor. Scholarly interest in spillover effects has grown markedly in recent years, and experimental designs to detect them have become increasingly sophisticated. Baird, McIntosh, and Ozler (2009) randomly varied the village-level saturation of cash transfers in Malawi to assess their effects on school attendance, and Sinclair, McConnell, and Green (2012) use a multi-level design to assess within-household and neighbor-to-neighbor spillover effects of direct mail on voting behavior. See Benjamin-Chung et al. (2018) for a recent review of applications in biomedical research and Halloran and Hudgens (2018) for a discussion of design considerations in regard to vaccine studies. We are

[1]On the other hand, the lack of media messages presents an opportunity, as potential audiences are often eager for entertainment, especially if it can be consumed at little or no cost. This point is especially important for video entertainment, which is beyond the means of most villagers and arguably more memorable and instructive than audio entertainment. Moreover, because these audiences have little exposure to conflicting media viewpoints, there is little risk that a message conveyed through education-entertainment will be undercut by other media sources, as often occurs in developed countries (McGuire, 1986; Chong and Druckman, 2013).

aware of no media experiments that assess spillover effects in a design-based fashion.

The current project represents an attempt to develop a research paradigm to assess the effectiveness of education-entertainment messaging on audiences as well as others in their social network. Our design makes use of the fact that Uganda, like much of sub-Saharan Africa, is dotted with video halls, small establishments where customers come to watch television and movies. By randomizing the content presented in these video halls and surveying those living nearby some weeks later, we can assess whether our messages have enduring effects on audiences and whether these effects diffuse to others in the community. This novel design also enables us to deploy messages on multiple topics and measure the effects of each message on beliefs, attitudes, and behaviors relevant to that topic. The use of multiple treatments and outcomes greatly augments our power to detect media effects, be they direct or second-hand.

The paper is structured as follows. We begin by providing an overview of our experimental design. Next, we summarize the videos we deployed in two rounds of experiments in different regions of Uganda, one conducted in 56 rural villages in 2015 and another in 112 rural villages in 2016. The third section describes the sampling procedures and outcome measures used in follow-up surveys conducted months after the media interventions ended. After explicating and verifying the core statistical assumptions underlying our design and analysis, we present our results.

We find substantial and highly significant direct effects in each of the three issue domains, teacher absenteeism, abortion stigma, and violence against women (VAW). The outcomes that reveal sizable direct effects, however, show faint spillover effects on others in the community, despite the fact that our design is well-powered to detect spillovers. This conclusion holds up even when we partition respondents by gender, as specified in our pre-analysis plan. Although direct effects often vary for male and female audience members, and although friendship networks would seem to facilitate within-gender communication, analyzing the results separately for men and women reveals at most equivocal evidence of spillovers. That is the case even though the context of our study seems particularly conducive to spillovers. We target an audience that has little access to locally produced media content of high production value, increasing the likelihood that audience members will engage with our video material and share their viewing experience with others. Our sample also consists of rural communities that tend to be close-knit, which should facilitate the flow of information through social networks. We conclude that the effects of education-entertainment interventions may stem primarily from direct exposure rather than diffusion.

2 Design Overview

The two experiments featured the same basic design, so for ease of exposition, we highlight the key features that they share in common before describing each of the studies in detail in subsequent sections. The key components of the experimental design are site selection, the random assignment procedures, the manner in which we address self-selection in exposure to the media messages, features of the sampling design that contribute to unobtrusiveness, assessment of gradations of exposure via social networks, and the use of multiple treatments and outcome measures.

Site selection. In order to minimize the risk that subjects would be exposed to media messages other than the ones to which they were randomly assigned, we began by selecting villages that would be at least 4 kilometers apart from one another. Given the limited road network separating villages, this walking distance effectively prevents people from attending media presentations in other villages.

Random assignment. To make the viewing experience as naturalistic as possible and to maximize the effect of messages about social norms, we sought to expose audiences to the media messages as groups rather than as isolated individuals.[2] We randomized at the level of the village, attracting audiences to a centrally-located video hall. This method of clustered assignment potentially sacrifices statistical power because all members of a given community are assigned together to the same experimental condition. We blocked on available covariates in order to minimize between-cluster variability and thereby the loss of power due to clustering. By allocating treatments over a large number of clusters we are able to estimate the clustered standard errors in a reliable manner. In the end, clustered assignment resulted in only a small loss of precision.

Intervention. Our media messages were embedded in a free film festival that was held over a series of weekends. The films were popular American movies that were narrated by a well-known Luganda-speaking celebrity. We inserted our messages – vignettes described in detail below dramatizing teacher absenteeism, violence against women, and abortion stigma – into commercial breaks. In contrast to the films, our videos were written and acted by Ugandans, and the dialogue was in Luganda. Because audiences were attracted by the feature-length films, the "Compliers" who attended the screenings were expected to have the same background attributes, regardless of the experimental condition to which their village was assigned. We

[2] In other locations, we conducted lab-like experiments in which we presented the videos to individual villagers who viewed them on a tablet computer and listened on headphones.

validate this expectation in section C of the appendix.

Multiple Message and Outcome Design. In order to increase our power to detect media effects, our experiment in effect simultaneously evaluates three different sets of videos, one on each topic. Pre-testing of the survey indicated that rural Ugandans' attitudes on these three topics are virtually uncorrelated, and so we did not expect videos on one topic to influence opinions about a different topic. We randomly assigned villages to display videos on one, two, or none of the three topics (absenteeism, violence, abortion). These topics are broached in our post-intervention survey, which covers a wide array of topics for the sake of obscuring the connection between the survey and the viewing experience.

Sampling. In contrast to many media evaluations, we did not restrict our attention to those who viewed our media messages. We enumerated all households living within a predetermined radius of the video hall and drew a random sample of the residents. The resulting respondent pool therefore comprises viewers and non-viewers, which not only facilitates the study of spillovers, it has the further advantage of obscuring the connection between the media intervention and the survey, which began approximately two months after the media messages were aired.

Measurement of principal strata. In order to avoid priming respondents to think about the videos when answering our outcome measures, questions about attendance were asked at the very end of the survey conducted two months after the end of the film festival. Drawing on the framework laid out in our pre-analysis plan, we differentiate among four different latent groups within our subject pool. The first is composed of Compliers, those who attended at least one film (and were therefore exposed to the assigned video messages at least once). The second group comprises Indirect Compliers, those who did not attend the films but whose family members attended at least one screening. In the third group are Apprised Never-Takers, those who neither attended nor had family members who attended, but who were nevertheless aware of the film festival. Finally, Never-Takers are the residual group who were unaware of the screenings. Below, we discuss the assumptions required for estimating causal effects within each principal stratum and consider an alternative classification that requires weaker assumptions.

Fortunately, in both studies the response rates are very high, well over 90 percent, and item-level missingness is low. Thus, we have an unusually precise accounting of the relative sizes of the four strata. Pooling the two studies, 1,492 (19 percent) of 7,965 subjects were Compliers. Another 3,285 (41 percent) were friends or family members of Compliers. A further 1,484 (19 percent) reported knowing about the film festival, even if they did not have friends or family members in attendance. The remainder of the subjects were unaware of the festival. These proportions give a sense of how large spillover effects would have to be in order to have the same overall impact as direct effects on viewers.

Pre-registered analyses. In sum, our aim was to create and deploy dramatizations that would change audiences' opinions on the three topics. Our pre-analysis plan specified that in the event that we found significant evidence of opinion change among Compliers, we would investigate spillover effects among others in the community. It also specified that we would look not only at effects for the sample as a whole but also broken down by respondent gender.

3 Description of Film Festivals and Videos

Our intervention consists of nine short video vignettes – three for each of our issue areas – which were interspersed throughout movies screened in 56 and 112 different villages in round 1 and round 2, respectively. The videos on teacher absenteeism and abortion stigma were identical across the two festivals. Prior to the launch of the second study, the videos on violence against women were re-written and re-shot after round 1 found them to be ineffective. We here focus on the round 2 videos on violence against women, which produced significant treatment effects among Compliers.

The video vignettes are each between three and eight minutes long. They were produced in Luganda (the main language spoken in the Central Region of Uganda) using Ugandan actors so as to make it easy for Ugandan villagers to identify with the characters in the videos. While an overarching narrative runs through the three vignettes for a given topic area, each vignette can also be understood as a self-contained story in isolation from the other two. The videos can be viewed at this address: http://tiny.cc/uganda_media.

We inserted the video vignettes as intermissions within films that we screened free of charge during multi-week film festivals in our study villages. In round 1, the film festivals took place on four consecutive weekends from November to December of 2015. In round 2, the festivals comprised six films shown one per week over consecutive weekends, from July 30 to September 4, 2016. Figures 1 and 2 give an overview of the study timelines, including the data collection strategy that we describe in more detail below. We advertised the film festivals throughout the village using posters, flyers and, if available, public loudspeakers. The films were unrelated to the treatment messages interspersed through them.

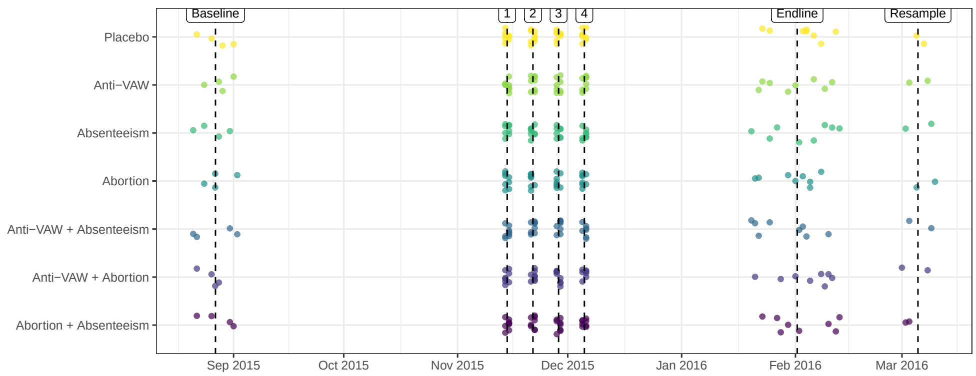

Figure 1: Timeline of round 1 media campaign and surveys.
Points represent unique visits to villages, either to screen films or to collect data. Colors and the Y axis represent the different treatment conditions. The X axis is ordered by date.

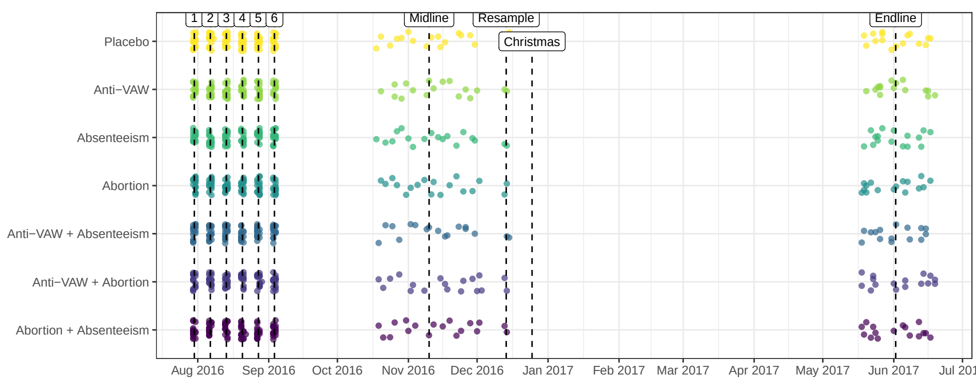

Figure 2: Timeline of round 2 media campaign and surveys.
Points represent unique visits to villages, either to screen films or to collect data. Colors and the Y axis represent the different treatment conditions, the X axis is ordered by date.

For each issue area, our video messages implicitly and explicitly express a prescriptive social norm.[3] For example, the videos emphasize the importance of speaking out against violence against women when it occurs by reporting cases to friends and family or village-level authorities. For abortion, the vignettes convey the obligation to help those who suffer from complications resulting from an abortion, irrespective of one's personal views about it. Finally, for teacher absenteeism, the videos convey the message that parents have a responsibility to take action in order to resolve the problem of teacher absenteeism. Additionally, the vignettes on all three topics convey the norm implicitly by modeling behavior in accordance with it. Each of the video

[3]Social psychologists use the term prescriptive social norm to refer to beliefs, statements or unwritten rules about appropriate behavior in a given situation (Shaffer, 1983).

vignettes ends with a narrator expressing the norm in a statement such as, "if you see domestic violence in your community, intervene before it's too late," "do not judge, and you will not be judged," or "we have a part to play in our children's education." A detailed description of the vignettes can be found in section A of the appendix.

Given their closeness to the audiences' context and experience, our video dramatizations seem well suited to make viewers identify with the main characters. It is rare for media with very high production value to be filmed in the local language (Luganda) using rural Ugandan villages as a setting. The videos depict situations that would be familiar to the participants in our study. And indeed the relevance of the films was apparent in a separate survey experiment we conducted wherein respondents were directly exposed to our video material on hand-held tablets prior to answering survey questions.[4] For example, after viewing the videos on domestic violence, the vast majority (84%) of respondents said that the stories could have happened in their village. That viewers found the stories relevant to their own lives is also reflected in what they said when invited to comment on the videos, for instance: "The video is so real" or "What I have seen in the video can also happen in my home."

4 Sampling

In both rounds we selected the villages in which to conduct our study using a non-random procedure designed to minimize spillovers between clusters and maximize statistical efficiency. A detailed discussion of the cluster sampling strategy can be found in section B of the appendix.

We conducted endline surveys that began roughly two months after the end of the film festivals.[5] We sampled individual respondents from a circular area around the video hall that was used to screen the treatment messages. Enumerators received a map for each village that depicted a circle around this video hall with a radius of between 200 and 800 meters. The radius was chosen based on the population density of the given village as judged from satellite images. Enumerators worked with village leaders (LC1 chairpersons) or other knowledgeable members of the community to produce a list of all households that reside within the circle indicated on the map. From this list, 40 households were randomly selected in round 1. In round 2, we randomly selected 50 households from the list. Among the selected households, twenty (round 1) or twenty-five (round 2) were randomly chosen as households within which a female respondent

[4]We did not include questions about respondents' views on the videos in our main survey to preserve, as much as possible, its unobtrusive character.

[5]In round 1, we also implemented a baseline survey. Following a Solomon four-group approach the baseline was only conducted among a randomly-selected subset of the villages. In round 2, we have two rounds of endline measurements collected roughly two and eights months after the end of the film festival. To ensure comparability with round 1, we only make us of measurements from the two-months endline here.

would be interviewed by a female interviewer; in the remaining households men were interviewed by male enumerators.

Upon meeting each household, enumerators listed all adult household members (aged 18-65) of the relevant gender and randomly selected one of them as the respondent. If no respondent of the relevant gender resided in the selected household, another household was randomly chosen from the list of households within the circle around the video hall. Enumerators always interviewed respondents of the same gender. If a respondent could not be found during the first visit of the enumerators, two additional visits were conducted before the respondent was coded as a non-response.

In the second round, there was a slight change in the sampling strategy for adults after the survey had been completed in all villages belonging to the first block. Specifically, we narrowed the age range of adult respondents from 18-65 to 18-50 and increased the number of respondents per village from 40 to 50. The first change was made to oversample Compliers and the second was due to additional capacities in our survey team that we had not anticipated. Since the same sampling strategy was used among villages within the same block, there is no correlation between the sampling strategy and treatment assignment within block.

Preliminary analyses that we conducted after having completed the endline survey in the first round showed that some cluster-level samples had very few responses from adult respondents who had attended at least one film. Consequently, we undertook a second round of sampling to target such Compliers, aiming to survey 15 additional adult respondents in 14 clusters. To select the 14 clusters, we identified the two clusters in each of the 7 treatment conditions with the fewest Compliers.[6] We conducted this sampling by continuing the same random sequence of households generated in the endline, so that the sampled units are the same units that would have been sampled had we continued endline data collection. In order to over-sample Compliers, the sampling strategy within households was altered to target respondents between 18 - 35, aiming for a target of 2/3 men. This change in plan was reflected in an addendum to the pre-analysis plan submitted prior to revealing the second round of data collection.[7] In the second round, we pre-specified and followed the same procedure.

Figures 3 and 4 summarize the individual-level samples from round 1 and 2. In round 1, our response rate among adult respondents in the endline survey (main data collection and re-sampling combined) is 99%.

[6] If there were more clusters with the same number of Compliers, we randomly selected one among them.

[7] The original Pre-Analysis Plan and addendum may be found at http://egap.org/registration/2207.

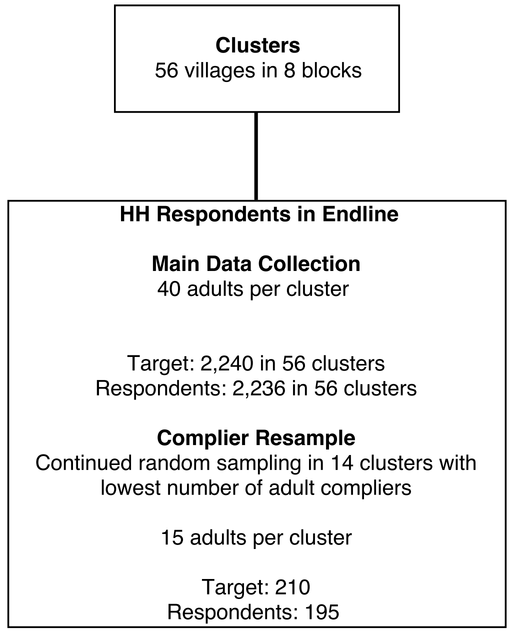

Figure 3: Individual-level sample round 1

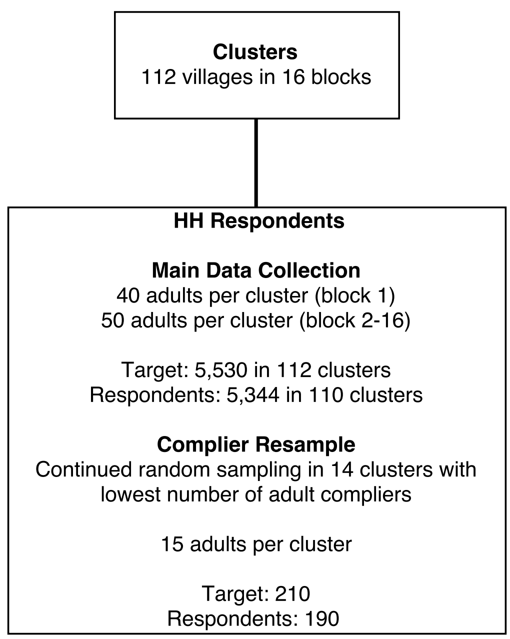

Figure 4: Individual-level sample round 2

In round 2, we were unable to conduct our household survey in two villages due to resistance from the local communities. We believe that our inability to work in these locations was unrelated to the treatment status of the village. The two locations are in an area known for suspicion towards outside groups. In both locations villagers were suspicious of the research team and in particular their motives for collecting head of household names (a component of the sampling procedure). There were fears related to land evictions and kidnapping. We deemed it unsafe to continue data collection in those areas. There was no indication from discussions with the residents of these villages that these difficulties were related to the specific treatment messages that were screened. In terms of the analysis, the above implies that we can recover an unbiased estimate of the average treatment effect among Compliers in villages that did not attrit. Therefore, in our main analysis we simply exclude villages in which we could not survey. Our response rate in round 2 is 96.4%.

In most of our analyses, we pool survey data from rounds 1 and 2, treating the two rounds as one experiment. Analyses of outcomes related to violence against women draw on round 2 data only, since, as discussed above, the treatment videos screened in round 2 were different from those used in round 1.

4.1 Random Assignment of Treatment

We randomly assigned individual survey respondents to seven treatment conditions at the village cluster level, within each block. The blocking scheme minimized within-block variance on latitude and longitude. Randomization was carried out using a random number generator in `R`. Three of these conditions involved the screening of video vignettes on one of the three topics, i.e., vignettes on domestic violence only, vignettes on abortion only, or vignettes on teacher absenteeism only. Another three conditions involved the screening of vignettes on two out of the three topics, i.e., vignettes on domestic violence and abortion, vignettes on domestic violence and absenteeism, and vignettes on abortion and absenteeism. The final condition is a pure placebo condition in which the film festival consisted only of the screening of Hollywood movies without any vignettes. One advantage of including a pure control group is that its respondents enable us to establish that opinions toward our three focal issues are very weakly correlated, which lends credence to the idea that our design offers three distinct tests in different attitude domains.

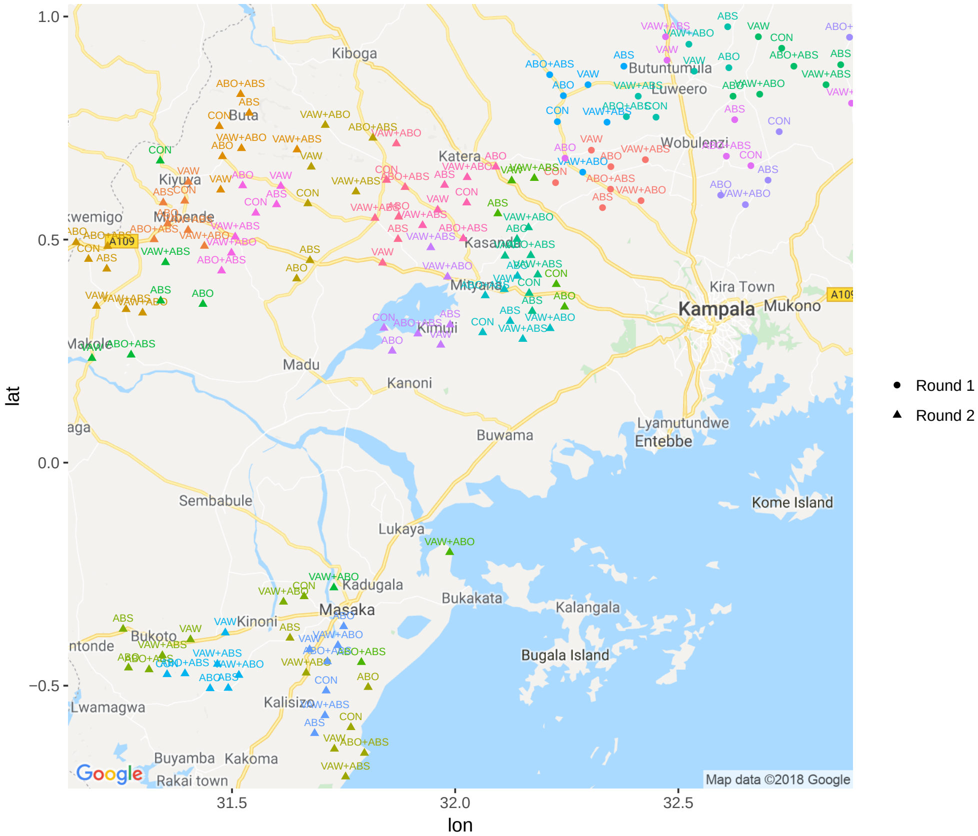

Figure 5: Clusters Included in the Study.

The map on figure 5 shows the location of the clusters in Round 1 and Round 2. The colors indicate the block, while labels indicate each of the seven treatment conditions: placebo (no vignettes), vignettes on abortion only, vignettes on domestic violence only, vignettes on teacher absenteeism only, vignettes on abortion and domestic violence, vignettes on abortion and absenteeism, and vignettes on domestic violence and absenteeism.

4.2 Compliance with Treatment at the Cluster Level

In Round 1, all clusters complied with the treatment assignment insofar as we were able to correctly screen the assigned films and messages in each village. In Round 2, two villages aired five of the six scheduled screenings. In one case, this was due to the video hall owner suspecting the feature-length film of spreading black magic; in another case, a local leader sought to prevent the screening apparently in an effort to extract a gratuity. In neither case do we have reason to suspect that this was due to the experimental vignette featured in the film. Because the extent of non-compliance is so minimal, we make no statistical correction for it.

5 Identification Strategy

5.1 Average Causal Effects for Each Latent Compliance Group

We denote a vector of random assignments, $\mathbf{z}^m$, where the superscript indicates the message to which the respondent was assigned,

$$m \in \{\text{placebo}, \text{VAW}, \text{abortion}, \text{absenteeism}\},$$

$z_i^m = 1$ when individual i in village j was assigned to message m, 0 otherwise. Note that this way of conceptualizing treatment assignment collapses across multiple of the seven treatment conditions described above. For example, $z_i^{\text{VAW}} = 1$ when individual i lives in a village that was assigned to messaging on violence against women, on violence against women and teacher absenteeism, or on violence against women and abortion. $z_i^{\text{VAW}} = 0$ when individual i resides in a village that was assigned to any of the other four treatment conditions.

At the individual level, we further define four types of compliance d_i with treatment assignment z_i^m, based on responses to the following two questions, which were asked at the end of our surveys (after all outcome measures). For example, in Round 2, we asked:

> Recently, a series of six free films (Pirates of the Caribbean, Creed, Fast and Furious, Spy, Slumdog Millionaire, Oz The Great And Powerful) were screened in the *kibanda* [video hall] in your village. Have you heard about the screenings and if so, how many screenings did you attend?
>
> Did your friends or family attend any of the screenings?

As summarized in table 1, we define individuals as being in direct compliance with their treatment assignment if they indicate that they attended at least one of the screenings. Indirect compliance means that individuals did not attend the film(s) themselves but report that family or friends attended. Those who report that they knew about the film(s) but did not attend any screenings and also do not report that friends and family attended are defined as being in apprised non-compliance with treatment assignment. We define individuals as not having complied if they report neither attending the screenings, knowing about the screenings, nor having direct family or friendship ties to those who attended.

Compliance Type	Answer to 19.1)	Answer to 19.2)
$d_i =$ direct compliance	1,2,3,4,5, or 6	anything
$d_i =$ indirect compliance	0 but knew about screenings 0 did not know about the screenings Don't know Refuse to answer	Yes, friends and family Yes, friends Yes, family
$d_i =$ apprised non-compliance	0 but knew about screenings	No Don't know Refuse to answer
$d_i =$ non-compliance	0 did not know about the screenings Don't know Refuse to answer	No Don't know Refuse to answer

Table 1: Definition of Compliance Types

We make the following identification assumption: $d_i(z_i^m = 1) = d_i(z_i^m = 0) = d_i$, for all m. In other words, the type of compliance is assumed to be a fixed personal attribute that is unaffected by treatment assignment. Given this assumption, we can denote four types of respondents:

$$s_i \in \{\text{Complier}, \text{Indirect Complier}, \text{Apprised Never-Taker}, \text{Never-Taker}\}.$$

In a placebo-controlled design such as ours, provided that this assumption is met, we obtain unbiased estimates of the average treatment affect among Compliers simply by comparing those who complied with the treatment (attended at least one film that contained a treatment video) to those who complied with the placebo (attended at least one film but never saw a treatment video). These direct effects may be characterized as Complier average causal effects, defined formally below. The same approach also applies to the spillover effects for Indirect Compliers, Apprised Never-Takers, and pure Never-Takers. For example, we compare Indirect Compliers in treatment locations to those in locations that did not air the treatment in order to assess the extent to which they were affect by second-hand exposure to the treatment by those in their social network.

The assumption that the treatment does not affect compliance is justified by the way in which the film festival was advertised and conducted. The focus was always on the films and not the messages conveyed during commercial breaks. One testable implication of this assumption is that the probability that a person engages in a a given form of compliance does not vary by treatment. We test and validate this stipulation in section D of the appendix. A further implication, corroborated in section C of the appendix, is that covariates are balanced among respondents in a given compliance stratum.

We are interested in the stratum-specific causal estimands, for example:

$$\tau^m_{\text{dir}} = E[(Y^m_i(Z^m = 1) - Y^m_i(z^m = 0) \mid s_i = \text{Complier}], \tag{1}$$

which reveals the average causal effect of the m messages on m-related outcomes among Compliers. Similarly,

$$\tau^m_{\text{ind}} = E[(Y^m_i(Z^m = 1) - Y^m_i(z^m = 0) \mid s_i = \text{Indirect Complier}], \tag{2}$$

$$\tau^m_{\text{app}} = E[(Y^m_i(Z^m = 1) - Y^m_i(z^m = 0) \mid s_i = \text{Apprised Never-Taker}], \tag{3}$$

$$\tau^m_{\text{nev}} = E[(Y^m_i(Z^m = 1) - Y^m_i(z^m = 0) \mid s_i = \text{Never-Taker}], \tag{4}$$

correspond to the average causal effect of messaging on m among Indirect Compliers (equation 2), Apprised Never-Takers (equation 3) and Never-Takers (equation 4), respectively. Our robustness check collapses Indirect Compliers, Apprised Never-Takers, and Never-Takers into a single group of Non-Compliers; this approach lacks nuance but guards against the possibility that the treatment affects how Non-Compliers are classified.

Our main analysis makes a second assumption to identify these quantities. Specifically, we presume that $Y^{m=k}(z^{m\neq k} = 1) = Y^{m=k}(z^{m=k} = 0)$, for all k. In other words, we assume no cross-over effects: m-specific outcomes are unaffected by assignment to non-m treatments. For example, we assume that violence-specific outcomes are unaffected by the absenteeism and abortion messages. In section E of the appendix, we show that our main results are robust to using an estimator that relaxes this assumption.

Given our identifying assumptions, we can estimate τ^m_{dir}, τ^m_{ind}, τ^m_{app} and τ^m_{nev} by fitting the following linear model among subsets of our data containing only one of the four s strata:

$$\mathbf{Y}^m = \gamma^m_0 + \mathbf{B}\boldsymbol{\gamma}^m + \tau^m_s \mathbf{z}^m + \mathbf{X}\boldsymbol{\delta}^m + \boldsymbol{\epsilon}^m. \tag{5}$$

Denoting by N_s the number of respondents in stratum s, $\mathbf{Y}^m$ is an N_s-length vector of observed outcomes related to message m, γ^m_0 is an intercept corresponding to the average of these outcomes among respondents in stratum s who were not assigned to message m in the reference block, $\mathbf{B}$ is an $N_s \times (K-1)$ matrix of block indicators and $\boldsymbol{\gamma}^m$ a $K-1$ vector of block fixed effects, τ^m_s is the effect of the m treatment in stratum s, $\mathbf{z}^m$ is an N_s-length vector indicating assignment to treatment m, $\mathbf{X}$ is an $N_s \times 2$ matrix of the average number of film festival attendees per village across all screenings and an indicator for whether a respondent was part of the complier resampling, $\boldsymbol{\delta}^m$ is a vector of corresponding effects, and $\boldsymbol{\epsilon}^m$ a vector of errors.

6 Outcome Measures

In order to work with more reliable outcome measures, we construct multi-item indexes from the myriad of survey responses to specific queries. Table 2 gives the wording and response distribution for each question that was used to construct a given index. The first index reflects what might be termed "conative attitudes" (Fishbein and Ajzen, 1975), or action orientations concerning teacher absenteeism. The common thread that runs through these questions is whether to address the problem of teacher absenteeism through collective action, as opposed to a passive approach, such as waiting for the situation to improve on its own. The Chronbach's alpha associated with this four-item index is 0.24; this relatively low value is not a source of bias but does make it especially challenging for us to detect treatment effects given the apparent signal to noise ratio. The second index draws on a much smaller pool of questions about the relative importance of educational goals such as reducing illiteracy or the "number of bad teachers at school." A third outcome measure concerns the topic of domestic violence. Again the common thread is respondents' conative attitudes toward helping a woman who has been beaten by her husband. Each respondent is presented with a series of paired options, one of which involves action (e.g., "I would accompany her to the police") while the other involves some kind of consolation that does not culminate in the involvement of authorities. The Chronbach's alpha associated with this four-item index is 0.4. Finally, we measure an outcome related to our abortion stigma messaging, intended to capture respondents' willingness to reach out to those facing stigma. We ask the respondent to imagine the case of a girl who has had an abortion and subsequently been ostracized from the community. If the respondent indicates that they would side with a friend who suggests the girl has "made her choice and has violated God's rule and we should not get involved", the response is coded as a 0. If instead the respondent agrees with a hypothetical friend who states "Regardless of what this girl did, we should be a friend to her and try to help her," we code the response 1.

Table 2: Outcome Measures

Index	Question	Value	Label	Round 1	Round 2
Act against absenteeism	Imagine that you find out that your child's teacher has been absent for 2 days this week during teaching hours. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take	0	Wait another few days to see if the problem corrects itself/Randomly assigned inaction	47%	34%
		1	Immediately begin organizing a PTA meeting, even if you know this might start some trouble	53%	66%
		NA	Don't know/Refuse	0%	0%
Act against absenteeism	Imagine that you find out that your child's teacher has been absent for 2 days this week during teaching hours. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take	0	Pray to god/Randomly assigned inaction	21%	48%
		1	Bring it up in the village meeting	79%	52%
		NA	Don't know/Refuse	0%	0%
Act against absenteeism	Imagine that you find out that your child's teacher has been absent for 2 days this week during teaching hours. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take	0	Send your child to a school in the neighboring village, where the teachers always come to class/Randomly assigned inaction	60%	68%
		1	Assemble a group of parents and confront the teacher	40%	32%
		NA	Don't know/Refuse	0%	0%
Act against absenteeism	Imagine that you find out that your child's teacher has been absent for 2 days this week during teaching hours. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take	0	Allow your child to leave school to help with the garden on days when the teacher is absent	2%	-
		1	Ask the headmaster to threaten to fire the teacher	98%	-
		NA	Don't know/Refuse	0%	-

Act against absenteeism	Imagine that you find out that your child's teacher has been absent for 2 days this week during teaching hours. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take	0	Randomly assigned inaction	-	25%
		1	Tell the LC1 chairperson to investigate why the headmaster has allowed this problem to occur	-	75%
		NA	Don't know/Refuse	-	0%
Education important goal	Here is a set of cards, which show different goals. Please choose the three that are the most important to you.	0	Did not choose 'Reducing the number of bad teachers at school'	58%	-
		1	Chose 'Reducing the number of bad teachers at school'	42%	-
		NA	Don't know/Refuse	0%	-
Education important goal	Here is a set of cards, which show different goals. Please choose the three that are the most important to you.	0	Did not choose 'Reducing illiteracy'	-	56%
		1	Chose 'Reducing illiteracy'	-	44%
		NA	Don't know/Refuse	-	0%
Act against VAW	Suppose you visit your cousin and she tells you that her husband beat her severely and asks you for help. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take.	0	Randomly assigned inaction	-	43%
		1	I would notify the Nabakyala and ask her to mediate the dispute	-	57%
		NA	Don't know/Refuse	-	0%
Act against VAW	Suppose you visit your cousin and she tells you that her husband beat her severely and asks you for help. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take.	0	Randomly assigned inaction	-	50%
		1	I would talk to her parents and ask them to come by to help the couple find a peaceful solution	-	50%
		NA	Don't know/Refuse	-	0%
Act against VAW	Suppose you visit your cousin and she tells you that her husband beat her severely and asks you for help. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take.	0	Randomly assigned inaction	-	81%
		1	I would accompany her to the police to report the incident	-	19%
		NA	Don't know/Refuse	-	0%

Act against VAW	Suppose you visit your cousin and she tells you that her husband beat her severely and asks you for help. Suppose there are only two actions that you can take. Please tell us which one you would prefer to take.	0	Randomly assigned inaction	-	72%
		1	I would get the LC1 chairperson involved	-	28%
		NA	Don't know/Refuse	-	0%
Help abortion	Suppose that a girl in your neighborhood has had a deliberate abortion. She has been ostracized from the community and people seem to have turned their backs on her. In this situation, two of your friends make the following two statements. Which friend would you agree with?	0	She made her choice and has violated god's rule and we should not get involved	27%	17%
		1	Regardless of what this girl did, we should be a friend to her and try to help her	73%	83%
		NA	Don't know/Refuse	0%	0%

Outcomes are combined into indices by averaging across them. Outcomes in the "Act against absenteeism" index ask respondents to choose one of two actions. In the first round, these two actions were fixed. In the second round, each "active" option was randomly paired with one of the following four "inactive" options: "Wait another few days to see if the problem corrects itself," "Send your child to a school in the neighboring village, where the teachers always come to class," "Find a tutor to instruct your child until the teacher comes back" or "Ask the headmaster to put your child into a different classroom until the teacher returns." For the outcomes in the "Act against VAW" index, the randomly assigned inaction was one of the following four options: "I would tell her that beating is often a sign of love and that she should try to work it out with her husband," "I would advise her to try harder to please her husband and things will likely improve," "I would express my sympathy for her but would tell her that every couple has to work it out for themselves" or "I would calm her down and tell her that the situation is bound to get better."

7 Results

We begin by considering the effects of messaging on teacher absenteeism on willingness to take action to counter absenteeism. The first column of Table 3 indicates that the mean for this outcome measure among the Compliers in the control condition is 0.61. The estimated treatment effect among the 1,492 Compliers is 0.040. The apparent increase in willingness to act is highly statistically significant ($p < 0.01$) and substantively large.[8] To put this estimate in perspective, the village-level standard deviation among Compliers in the control group is 0.11. Clearly, the messages concerning teacher absenteeism produced a statistically significant and meaningful shift in behavioral orientations among those who attended the screenings.

	Dependent variable:				
	Index of willingness to take action to counter absenteeism				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
absenteeism	0.040***	−0.001	0.015	−0.003	0.001
	(0.013)	(0.009)	(0.014)	(0.012)	(0.007)
Control Mean	0.61	0.59	0.59	0.6	0.59
Vill. Means	0.63	0.6	0.58	0.58	0.59
Vill. SD	0.11	0.08	0.12	0.11	0.07
N Vill.	166	166	166	165	166
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,492	3,285	1,484	1,704	6,473
Adjusted R^2	0.026	0.044	0.054	0.067	0.049
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 3: Direct effects and spillovers from absenteeism messages among all respondents in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

The apparent second-hand effects are more muted. Column 2 reports the estimated effect among 3,285 Indirect Compliers, whose family or friends attended the screenings. Treatment assignment had no apparent effect on Indirect Compliers, generating a weakly negative point estimate of -0.001 with a standard error of 0.009. Somewhat more convincing are the results for Apprised Never-Takers (N=1,484), for whom we estimate a second-hand treatment effect of 0.015 with a standard error of 0.014. The fraction of the direct effect that spilled over to

[8] Even though we had suggested one-tailed tests in our pre-analysis plan, it turned out that the choice between one-tailed and two-tailed tests is inconsequential. Here, we report two-tailed hypothesis tests.

Apprised Never-Takers is 0.015/0.050 or 30 percent, although this estimate falls well short of conventional levels of statistical significance. Reassuringly, we see no evidence whatsoever of spillovers among pure Never-Takers, whose point estimate is weakly negative. Pooling over all non-Compliers, we obtain a precise estimate that is very close to zero.

	Dependent variable:				
	Education is an important goal				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
absenteeism	0.056**	0.010	−0.035	0.001	−0.001
	(0.023)	(0.016)	(0.023)	(0.024)	(0.011)
Control Mean	0.44	0.43	0.45	0.42	0.43
Vill. Means	0.43	0.43	0.44	0.43	0.43
Vill. SD	0.19	0.12	0.2	0.19	0.08
N Vill.	166	166	166	165	166
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,492	3,285	1,484	1,704	6,473
Adjusted R^2	0.008	−0.001	0.0002	0.007	−0.0005
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 4: Direct effects and spillovers from absenteeism messages among all respondents in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

A further outcome of interest is whether respondents rate education-related goals among the most important goals for their village (see Table 4). Among Compliers in the control group, the mean is 0.44. The treatment effect among Compliers is again substantial, 0.056 with a standard error of 0.023 ($p < 0.05$). Among Indirect Compliers, we find a positive point estimate (0.010) that is eclipsed by its standard error (0.016). This estimate falls well short of statistical significance, but taken at face value it suggests that 0.01/0.056 or 18 percent of the direct treatment effect is transmitted to Indirect Compliers. This time, however, the estimated effect among Apprised Never-Takers is unexpectedly negative, which suggests that sampling variability may account for any apparent spillover effects detected earlier for this subgroup. Again, as expected, we see no evidence whatsoever of spillovers to pure Never-Takers. Pooling over all non-Compliers again renders an estimate very close to zero.

Turning to the issue of violence against women, Table 5 reports the results from regressions in which the outcome is willingness to take action to assist victims and report incidents to village authorities. The mean among Compliers in the control group on this index is 0.38, and the average treatment effect for this subgroup is estimated to be 0.047 with a standard error of 0.015 ($p < 0.01$). The point estimate again suggests a substantial effect in light of the fact that the village-level standard deviation is 0.09. Very little of this effect was transmitted to Indirect Compliers, for whom the point estimate is just 0.004 (SE=0.012). The point estimate for Apprised Never-Takers is weakly negative.

On the topic of abortion stigma, Table 6 indicates willingness to help those suffering from post-abortion complications increased significantly among Compliers. Estimated spillover effects are weakly negative for both Indirect Compliers and Apprised Never-Takers, and well short of statistical significance. Pooling over all non-Compliers, we obtain a weakly negative point estimate. In sum, we find little or no evidence of spillover effects.

	Dependent variable:				
	Index of willingness to take action to counter intimate partner violence				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
VAW	0.047***	0.004	−0.008	0.009	0.003
	(0.015)	(0.012)	(0.018)	(0.017)	(0.010)
Control Mean	0.38	0.38	0.38	0.37	0.38
Vill. Means	0.38	0.37	0.37	0.37	0.38
Vill. SD	0.09	0.07	0.12	0.12	0.06
N Vill.	110	110	110	109	110
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,156	2,447	953	978	4,378
Adjusted R^2	0.013	0.005	0.023	0.009	0.007
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 5: Direct effects and spillovers from anti-VAW messages among all respondents in endline surveys following 2016 festival.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:				
	Willingness to help someone suffering from post-abortion complications				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
abortion	0.043^{**}	-0.002	-0.030	0.017	-0.003
	(0.019)	(0.014)	(0.022)	(0.021)	(0.012)
Control Mean	0.82	0.8	0.8	0.78	0.79
Vill. Means	0.79	0.79	0.8	0.78	0.79
Vill. SD	0.17	0.13	0.16	0.18	0.11
N Vill.	166	166	166	165	166
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,492	3,285	1,484	1,704	6,473
Adjusted R^2	0.009	0.019	0.036	0.010	0.019
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 6: Direct effects and spillovers from anti-abortion stigma messages among all respondents in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

7.1 Effects by Gender

Because gender defines the lines of communication within villages and because treatment effects among Compliers may vary between men and women, we further investigate whether the apparent patterns of spillover effects change when we focus our attention solely on men or women. Table 7 suggests that with regard to taking action to address teacher absenteeism, Complier average causal effects appear to be somewhat larger for men (0.046, SE=0.015) than for women (0.017, SE=0.023), although the treatment-by-gender interaction is not significant. We do not find markedly greater spillovers for men than women, however. Among male Indirect Compliers the point estimate is weakly positive (0.003, SE=0.012), while for women it is weakly negative (-0.003, SE=0.013).

	Dependent variable:			
	Index of willingness to take action to counter absenteeism			
	Compliers - Men	Indirect Compliers - Men	Compliers - Women	Indirect Compliers - Women
	(1)	(2)	(3)	(4)
absenteeism	0.046^{***}	0.003	0.017	-0.003
	(0.015)	(0.012)	(0.023)	(0.013)
Control Mean	0.64	0.63	0.54	0.54
Vill. Means	0.65	0.64	0.56	0.55
Vill. SD	0.11	0.1	0.2	0.09
N Vill.	165	166	143	166
Block FE	Yes	Yes	Yes	Yes
Observations	1,017	1,716	475	1,569
Adjusted R^2	0.021	0.060	0.062	0.027
Notes:				*p<0.1; **p<0.05; ***p<0.01

Table 7: Direct effects and spillovers from absenteeism messages among men and women in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:			
	Education is an important goal			
	Compliers - Men	Indirect Compliers - Men	Compliers - Women	Indirect Compliers - Women
	(1)	(2)	(3)	(4)
absenteeism	0.040	-0.001	0.103^{**}	0.023
	(0.029)	(0.023)	(0.043)	(0.023)
Control Mean	0.48	0.45	0.36	0.41
Vill. Means	0.47	0.45	0.34	0.41
Vill. SD	0.23	0.15	0.33	0.19
N Vill.	165	166	143	166
Block FE	Yes	Yes	Yes	Yes
Observations	1,017	1,716	475	1,569
Adjusted R^2	0.0003	-0.004	0.013	0.008
Notes:				*p<0.1; **p<0.05; ***p<0.01

Table 8: Direct effects and spillovers from absenteeism messages among men and women in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:			
	Index of willingness to take action to counter intimate partner violence			
	Compliers - Men	Indirect Compliers - Men	Compliers - Women	Indirect Compliers - Women
	(1)	(2)	(3)	(4)
VAW	0.028	−0.002	0.104***	0.010
	(0.019)	(0.014)	(0.027)	(0.018)
Control Mean	0.4	0.38	0.35	0.37
Vill. Means	0.4	0.38	0.35	0.37
Vill. SD	0.12	0.09	0.19	0.1
N Vill.	110	110	97	110
Block FE	Yes	Yes	Yes	Yes
Observations	797	1,253	359	1,194
Adjusted R^2	−0.002	0.010	0.070	−0.001
Notes:				*p<0.1; **p<0.05; ***p<0.01

Table 9: Direct effects and spillovers from anti-VAW messages among men and women in endline surveys following 2016 festival.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:			
	Willingness to help someone suffering from post-abortion complications			
	Compliers - Men	Indirect Compliers - Men	Compliers - Women	Indirect Compliers - Women
	(1)	(2)	(3)	(4)
abortion	0.012	−0.003	0.124***	−0.007
	(0.020)	(0.017)	(0.040)	(0.022)
Control Mean	0.86	0.84	0.73	0.75
Vill. Means	0.84	0.82	0.67	0.74
Vill. SD	0.18	0.15	0.36	0.2
N Vill.	165	166	143	166
Block FE	Yes	Yes	Yes	Yes
Observations	1,017	1,716	475	1,569
Adjusted R^2	−0.00002	0.022	0.043	0.026
Notes:				*p<0.1; **p<0.05; ***p<0.01

Table 10: Direct effects and spillovers from anti-abortion stigma messages among men and women in endline surveys following 2015 and 2016 festivals.

Coefficients estimated using the pre-registered least-squares regression, conditioning on block fixed-effects and an indicator for resampling. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

Women, on the other hand, seem to be more responsive to the message that improving education is an important community goal. The direct effect of exposure for female Compliers is 0.103 (SE=0.043), as compared to 0.040 (SE=0.029) for male Compliers. The corresponding spillover effects for Indirect Compliers follow the hypothesized pattern, but only to a limited extent: for women, the estimated effect is 0.023 (SE=0.023), whereas for men it is weakly negative (-0.001, SE=0.023).

Whereas for absenteeism the treatment-by-gender interaction is insignificant for Compliers, this interaction is significant ($p < 0.05$) for the topic of domestic violence. Female Compliers are highly responsive to the videos, with an estimated average effect of 0.104 (SE=0.027), as compared to men, among whom opinion change is fairly limited (0.028, SE=0.019). However, we find relatively little evidence of spillover effects to Indirect Compliers. Among men, the point estimate is weakly negative, and among women it is just 0.010 (SE=0.018), implying that less than 10 percent of the direct effect is transmitted to Indirect Compliers.

Finally, we again see a significant treatment-by-gender interaction for willingness to help those suffering from post-abortion complications, with female Compliers showing quite strong effects. Yet, the estimated spillover effect for Indirect Compliers is weakly negative.

Overall, the analysis by gender offers little support for the hypothesis that treatment effects spill over from audience members to others of the same sex. Even focusing on instances in which the estimated direct effect of the messages is especially strong for one sex or the other, we nonetheless find limited evidence of transmission to Indirect Compliers.[9]

8 Conclusion

This paper makes two contributions to the study of media effects in developing countries. The first is methodological. We propose a placebo-controlled design to assess the effects of media messages on audiences and others in their social networks. The design allows for unobtrusive assessment of the extent to which media messages change beliefs, attitudes, and behaviors among different segments of the target population. We implemented this design in two successive experiments involving thousands of villagers in more than 150 villages. Although the design imposes a set of important assumptions about the comparability of audiences and their social networks across treatment conditions, it also allows for diagnostic tests of these assumptions, and both studies seem to have satisfied the requirements of the experimental design.

[9] In section F of the appendix, we make use of the fact that our design allows for the simultaneous test of messages on a variety of topics. We conduct a broader and more powerful test of the overall influence of media messaging, both on Compliers and various gradations of non-Compliers. In line with the findings presented here, the results suggest that spillovers, if they do occur, are small.

The second contribution is substantive. The literature on media effects has speculated about the possibility of multiplier effects due to communication between audiences and others in their social network. Because the number of audience members tends to be considerably smaller than the number of people in their social networks, even relatively small average spillover effects may imply cumulative effects that rival the effects of direct exposure. Our study took place in a setting that seems particularly prone to such second-hand effects: The education-entertainment messages feature local actors speaking the local language and provided an immersive experience to an audience with limited access to visual media. Moreover, we worked in relatively small and tight-knit communities: 50% of the respondents in our second-round sample indicated that they discuss things that are going on in the village every day or almost every day with nearby-neighbors and 54% of respondents in our first-round sample say that they would be able to name everyone or almost everyone in their village. The results from our experiments suggest that even in instances where diffusion appears likely and where direct effects on audiences' opinions are large and statistically significant, second-hand effects seem meager, whether considered individually or jointly.

Because this is the first study of its kind, the open question is whether the lack of spillover effects is a general feature of dramatized messages or rather specific to the issues or format of the videos used here. Our messages were interspersed in a larger feature-length film, and it remains to be seen whether spillover effects are more evident when the experimental treatment is the main event rather than a side show. Our videos also depict social problems and their tragic consequences; one wonders whether such "heavy" storylines discourage the kinds of interpersonal conversation through which media effects may be transmitted. Much work remains to be done to develop an evidence-based understanding of the conditions under which spillovers occur. Until then, those who seek to influence attitudes and behavior via dramatized messages should focus primarily on enlarging the audiences who are directly exposed.

References

Abramsky, Tanya, Karen Devries, Ligia Kiss, Janet Nakuti, Nambusi Kyegombe, Elizabeth Starmann, Bonnie Cundill, Leilani Francisco, Dan Kaye, Tina Musuya, Lori Michau, and Charlotte Watts. 2014. "Findings from the SASA! Study: a cluster randomized controlled trial to assess the impact of a community mobilization intervention to prevent violence against women and reduce HIV risk in Kampala, Uganda." *BMC Medicine* 12: 122–139.

Babalola, Stella, Angela Brasington, Ada Agbasimalo, Anna Helland, Edith Nwanguma, and Nkechi Onah. 2006. "Impact of a communication programme on female genital cutting in eastern Nigeria." *Tropical Medicine & International Health* 11 (10): 1594–1603.

Bandura, Albert. 2004. "Social Cognitive Theory for Personal and Social Change by Enabling Media." In *Entertainment-Education and Social Change: History, Research, and Practice*, ed. Arvind Singhal, Michael J. Cody, Everett M. Rogers, and Miguel Sabido. Mahwah, New Jersey: Lawrence Erlbaum pp. 75–96.

Banerjee, Abhijit, Sharon Barnhardt, and Esther Duflo. 2017. "Movies, Margins and Marketing: Encouraging the Adoption of Iron-Fortified Salt." In *Insights in the Economics of Aging*, ed. David A. Wise. Chicago and London: University of Chicago Press pp. 285–306.

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.

Blair, Graeme, Rebecca Littman, and Elizabeth Levy Paluck. 2017. "Motivating the Adoption of New Community-Minded Behaviors: An Empirical Test in Nigeria." Unpublished manuscript. **URL:** *https://papers.ssrn.com/sol3/papers.cfm?abstract_ id=3033133*

Bowers, Jake, Mark M Fredrickson, and Peter M Aronow. 2016. "Research note: a more powerful test statistic for reasoning about interference between units." *Political Analysis* 24 (3): 395–403.

Chong, Dennis, and James N Druckman. 2013. "Counterframing effects." *The Journal of Politics* 75 (01): 1–16.

Fishbein, Martin, and Icek Ajzen. 1975. *Belief, attitude, intention, and behavior: An introduction to theory and research.* MA: Addison-Wesley.

Galiani, Sebastian, Paul Gertler, and Alexandra Orsola-Vidal. 2012. "Promoting Handwashing Behavior in Peru: The Effect of Large-Scale Mass-Media and Community Level Interventions." World Bank Policy Research Working Paper No. 6257.
URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2170640*

Green, Donald, and Srinivasan Vasudevan. 2015. "Diminishing the Effect of Vote-buying on Electoral Outcomes in India: A Pilot RCT to Test the Effectiveness of Radio Messages." *Working Paper* .

Halloran, M Elizabeth, and Michael G Hudgens. 2018. "Estimating population effects of vaccination using large, routinely collected data." *Statistics in medicine* 37 (2): 294–301.

Heatherton, Todd F, and James D Sargent. 2009. "Does watching smoking in movies promote teenage smoking?" *Current Directions in Psychological Science* 18 (2): 63–67.

Heong, KL, MM Escalada, NH Huan, VH Ky Ba, PV Quynh, LV Thiet, and HV Chien. 2008. "Entertainment–education and rice pest management: A radio soap opera in Vietnam." *Crop Protection* 27 (10): 1392–1397.

Jensen, Robert, and Emily Oster. 2009. "The power of TV: Cable television and women's status in India." *The Quarterly Journal of Economics* pp. 1057–1094.

La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea. 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics* pp. 1–31.

McGuire, William J. 1986. "The myth of massive media impact: Savagings and salvagings." *Public communication and behavior* 1: 173–257.

Paluck, Elizabeth Levy. 2009. "Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda." *Journal of personality and social psychology* 96 (3): 574–587.

Paluck, Elizabeth Levy. 2010. "Is it better not to talk? Group polarization, extended contact, and perspective taking in eastern Democratic Republic of Congo." *Personality and Social Psychology Bulletin* 36: 1170–1185.

Paluck, Elizabeth Levy, and Donald P. Green. 2009. "Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in Rwanda." *American Political Science Review* 103 (4): 622–644.

Petty, Richard, and John T Cacioppo. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York: Springer.

Rosenbaum, Paul R. 2002. *Observational studies.* New York: Springer.

Sarrassat, Sophie, Nicolas Meda, Moctar Ouedraogo, Henri Some, Robert Bambara, Roy Head, Joanna Murray, Pieter Remes, and Simon Cousens. 2015. “Behavior change after 20 months of a radio campaign addressing key lifesaving family behaviors for child survival: midline results from a cluster randomized trial in rural Burkina Faso.” *Global Health: Science and Practice* 3 (4): 557–576.

Shaffer, Leigh S. 1983. “Toward Pepitone’s vision of a normative social psychology: What is a social norm?” *The journal of mind and behavior* pp. 275–293.

Singhal, Arvind, Everett M Rogers, and William J Brown. 1993. “Harnessing the potential of entertainment-education telenovelas.” *International Communication Gazette* 51 (1): 1–18.

UNFPA-UNICEF. 2014. “Voices of Change.” *Annual Report on Joint Programme on FGM/Cutting: Accelerating Change* .

UNICEF. 2005. “Violence against Disabled Children.” *UN Secretary Generals Report on Violence against Children Thematic Group on Violence against Disabled Children* .

Appendix to:

A Placebo Design to Detect Spillovers from an Education-Entertainment Experiment in Uganda

A Detailed Description of Video Vignettes

For teacher absenteeism, the videos show that acting in accordance with the norm is effective in bringing absent teachers back to the classroom. Unlike domestic violence and abortion, teacher absenteeism is uncontroversially viewed as a social bad. The vignettes depict parents who, upon learning their children's teacher has not been coming to class, organize a meeting among members of the parent teacher association (PTA). In the first vignette they discover that the teacher has been absent from classes. In the second, they learn that the teacher has not been paid and is selling soap in the market in order to make ends meet. In the final vignette, the parents' action results in the school being monitored by a government official, whose oversight of the headmaster forces him to pay his teachers. Throughout the story, parents emphasize their responsibility to ensure that their children receive the good education that they deserve.

The vignettes on violence against women (revised for the second round study in 112 villages) contrast a storyline in which a victim of violence does not receive help with one in which the community steps in to ameliorate the situation. In the first vignette the protagonist is a sympathetic and personable woman whose husband beats her severely despite her sincere efforts to appease him. The protagonist's neighbor overhears her screams but decides not to speak out. In the second vignette, which begins with the protagonist's hospitalization and ends with her funeral, we learn that not only her neighbor, but also her daughter and parents knew about the violence. They express regret for failing to speak out sooner. In the third vignette, the setting is a nearby village in which a similar story is unfolding. The focal woman in the story is also beaten by her husband, but unlike the woman in the preceding vignette, she decides to disclose this information to her parents. Rather than scold, her parents intervene to help mediate. Moreover, the parents share the information with the local women's counselor (*Nabakyala*), who visits the household to provide guidance. The vignette closes with the couple in visibly better relations with one another.

In the case of abortion, the vignettes are designed to inspire empathy with women who are stigmatized as a result of having chosen to have an abortion. The videos follow a young woman who suffers from painful and debilitating post-abortion complications. Upon being discovered by her aunt in the first vignette, she confesses that she kept quiet about the abortion and her symptoms for fear of being ostracized by her friends and family. The aunt voices compassion and rushes her niece to the hospital, where, in the second vignette, the aunt engages in a conversation with the doctor. The doctor reiterates the obligation to help those suffering from the consequences of botched abortions and speaks out against harsh judgment. He explains that abortions are a common occurrence and warns against the medical risks of late or no treatment, connecting them to the fear of ostracism. In the final vignette, a school teacher reassures the now-recovered girl that she can come back to school without fear of being expelled.

B Sampling Strategy for Clusters

The first round of our experiment took place in 56 villages in the districts of Luwero, Kayunga and Nakeseke. The second round took place in 112 villages in the districts of Mubende, Mityana, Masaka and Lwengo. To select these villages, we first identified villages with video halls in the relevant districts. This led to a set of 162 candidate villages in round 1 and approximately 300 candidate villages in round 2. We then identified and excluded potentially problematic sites (e.g., video halls that operated seasonally or did not have backup generators), narrowing down the selection to 123 candidate villages in round 1 and 247 candidate villages in round 2.

In order to minimize spillovers between villages, we used a random walk procedure to identify subsets of candidate villages in which all members were at least some specific euclidean distance from one another. In round 1, we permuted one million unique sets of 64 villages from the candidate villages, such that each village was at least 5 kilometers from its closest neighbor. In round 2, we chose sets of 125 villages with a distance constraint of 4.5 kilometers.

From among these subsets that met our distance constraint, we selected the sets of villages that maximized the total distance between villages, as well as the number of video-halls within villages in the set (some villages have more than one video hall).

In round 1, we employed a blocking algorithm to organize the sets of 64 villages into 8 blocks of 8 villages, minimizing within-block variance on latitude, longitude, and the approximate population size of each village (coded by workers on Amazon Mechanical Turk using satellite imagery). We selected the set of 64 villages that minimized the total within-block variance, and non-randomly excluded one village from each block, giving a sample of 56 villages in eight blocks (all prior to randomization).

In round 2, we only relied on the total distance between villages to make the selection of distance-eligible subsets. Due to logistical constraints, we replaced 19 villages in the initially selected set by hand-selecting other clusters sufficiently distant from the remaining set. Among the eligible set of 125 villages, we chose the 112 villages with the largest distance to the nearest unit. See figures 6 and 7 in the appendix for a graphical depiction of the cluster sampling.

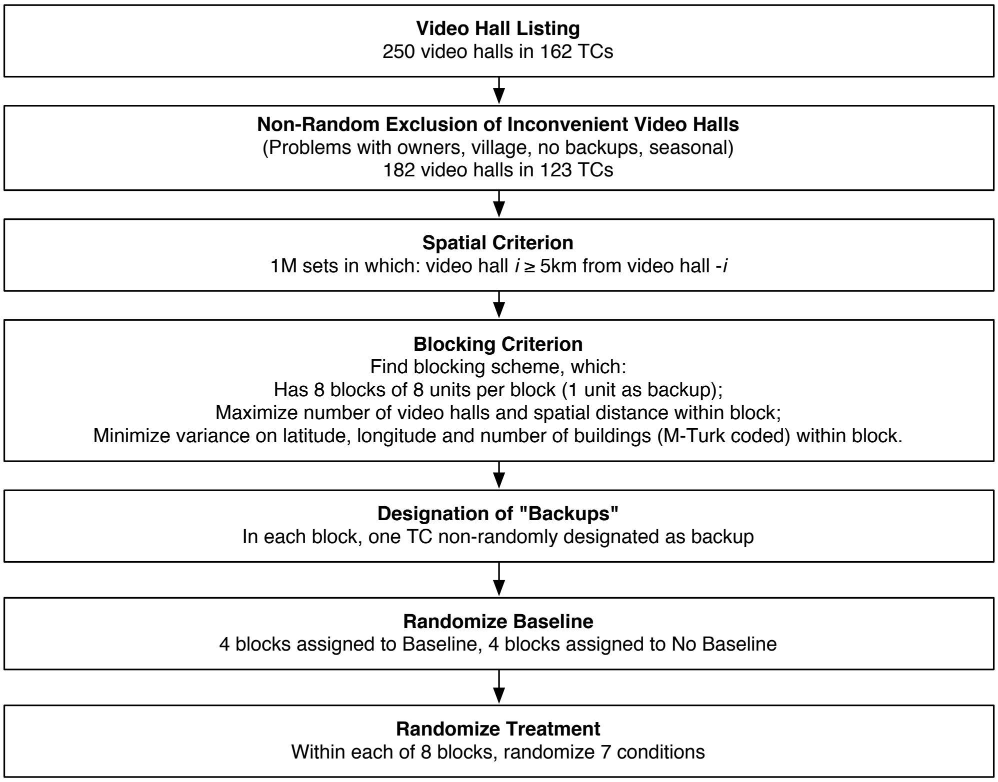

Figure 6: Sampling of clusters in round 1

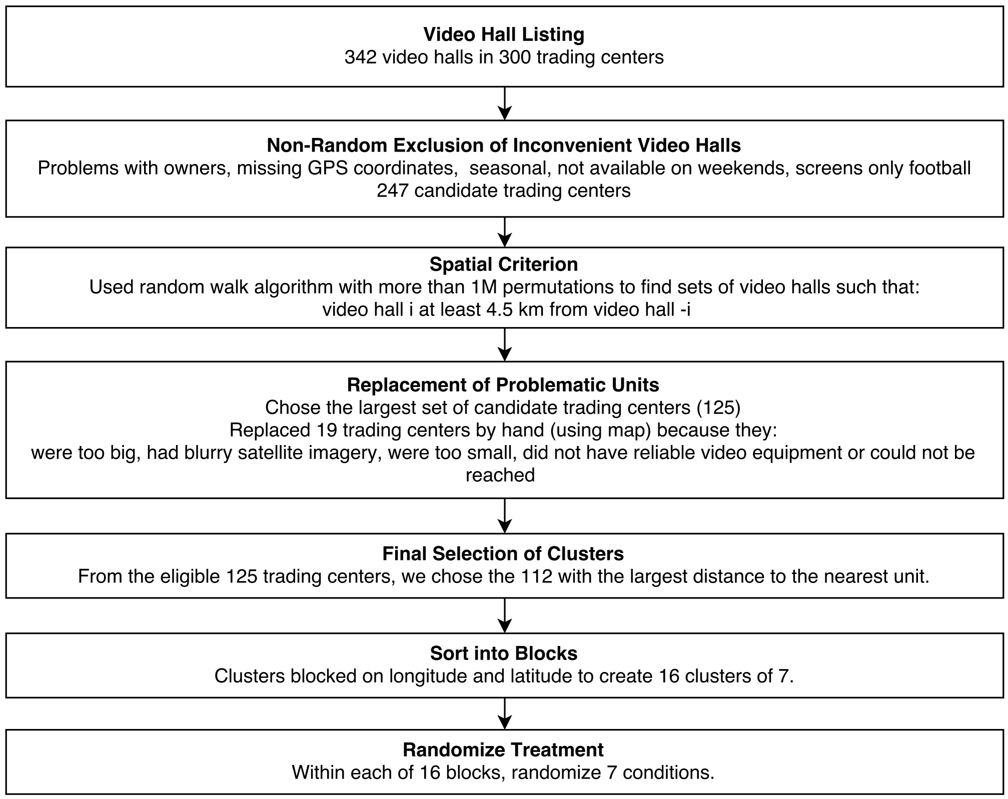

Figure 7: Sampling of clusters in round 2

C Balance on Covariates among Compliance Strata

We examine balance on observable pre-treatment covariates for Compliers, Indirect Compliers, Apprised Never-Takers, and Never-Takers. For each covariate, we test for a significant relationship to the treatment using randomization inference to conduct a likelihood ratio test. In the tables below, the first column names the covariate and the following seven columns show means of covariate under the respective treatment conditions. The last column in the table shows the p-value from the likelihood ratio test. The 'full' model regresses the covariate on the six non-placebo treatment indicators, controlling for block and resample fixed effects. The restricted model regresses the covariate on block and resample fixed effects only. The observed likelihood ratio is compared to 2000 likelihood ratios simulated under the null of no effect of treatment on the covariate for all units by re-permuting the treatment assignment and re-estimating the likelihood. The p-value is equal to the proportion of such simulations at least as great as the observed likelihood ratio. Note that p-values are not adjusted to account for family-wise error rates: under independence, in expectation $x\%$ of the covariates should exhibit imbalance that is significant at the $x\%$ level. The balance tables can be summarized as follows:

- Table 11 reports balance of 83 covariates across the seven treatment conditions among all Compliers: 5/83 (6%) tests exhibit a p-value equal to or less than .05.
- Table 12 reports balance of 83 covariates across the seven treatment conditions among all Indirect Compliers: 3/83 (4%) tests exhibit a p-value equal to or less than .05.
- Table 13 reports balance of 83 covariates across the seven treatment conditions among all Apprised Never-Takers: 8/83 (10%) tests exhibit a p-value equal to or less than .05.
- Table 14 reports balance of 83 covariates across the seven treatment conditions among all pure Never-Takers: 4/83 (5%) tests exhibit a p-value equal to or less than .05.
- Table 15 reports balance of 83 covariates across the seven treatment conditions among all Non-Compliers (Indirect Compliers, Apprised Never-Takers and Never-Takers): 2/83 (2%) tests exhibit a p-value equal to or less than .05.

	PLA	VAW	ABO	ABS	ABO_ABS	VAW_ABS	VAW_ABO	p-value
protestant	0.24	0.28	0.14	0.13	0.18	0.16	0.19	0.00
english_christian	0.03	0.05	0.07	0.07	0.06	0.05	0.16	0.01
minority_lang	0.07	0.05	0.09	0.03	0.06	0.03	0.02	0.04
munyankole	0.06	0.08	0.08	0.04	0.08	0.16	0.11	0.05
domestic_work	0.02	0.05	0.07	0.03	0.06	0.02	0.04	0.05
other_person	0.04	0.08	0.09	0.12	0.06	0.03	0.09	0.07
minority_tribe	0.20	0.21	0.13	0.13	0.17	0.08	0.09	0.07
atheist	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.10

misc_floor	0.09	0.10	0.20	0.12	0.16	0.12	0.18	0.11
rooms	2.48	2.44	2.20	2.51	2.71	2.54	2.41	0.13
misc_light	0.12	0.14	0.18	0.09	0.17	0.12	0.12	0.14
munyarwanda	0.07	0.04	0.13	0.11	0.09	0.13	0.08	0.14
munyoro	0.08	0.04	0.05	0.08	0.03	0.03	0.07	0.15
highest_grade	7.28	7.77	6.67	7.06	7.20	6.53	7.08	0.16
write_only	0.05	0.03	0.06	0.04	0.02	0.04	0.02	0.17
holy_spirit	0.08	0.08	0.14	0.14	0.13	0.15	0.09	0.18
cellphone	0.85	0.75	0.75	0.85	0.79	0.79	0.78	0.23
luganda_lang	0.89	0.93	0.86	0.88	0.85	0.83	0.94	0.23
catholic	0.41	0.38	0.45	0.46	0.43	0.50	0.36	0.24
university	0.05	0.04	0.03	0.05	0.06	0.02	0.02	0.24
write_and_read	0.81	0.88	0.79	0.84	0.85	0.82	0.88	0.26
other_work	0.07	0.09	0.05	0.04	0.04	0.05	0.03	0.28
runyannkole_lang	0.02	0.02	0.04	0.02	0.04	0.11	0.04	0.29
chair	0.80	0.85	0.79	0.88	0.85	0.81	0.84	0.30
age	28.46	28.48	28.20	28.63	29.71	29.98	28.47	0.34
solar_light	0.27	0.29	0.28	0.39	0.34	0.32	0.24	0.36
female	0.25	0.32	0.35	0.28	0.33	0.38	0.30	0.36
education_work	0.06	0.06	0.02	0.03	0.02	0.04	0.05	0.39
mukiga	0.02	0.02	0.03	0.02	0.08	0.03	0.03	0.40
separated	0.11	0.13	0.12	0.09	0.09	0.14	0.11	0.45
read_only	0.04	0.02	0.04	0.03	0.02	0.01	0.02	0.45
misc_wall	0.03	0.04	0.07	0.01	0.03	0.06	0.06	0.46
manual_work	0.10	0.06	0.07	0.07	0.09	0.11	0.07	0.48
same_village	0.52	0.44	0.43	0.48	0.52	0.45	0.47	0.51
tv	0.26	0.16	0.19	0.25	0.18	0.17	0.19	0.52
not_married	0.23	0.24	0.18	0.22	0.18	0.25	0.21	0.52
transport_work	0.04	0.04	0.04	0.06	0.03	0.06	0.09	0.52
motor_cycle	0.22	0.22	0.22	0.27	0.24	0.20	0.19	0.54
mobile_phone_use	3.50	3.23	3.14	3.37	3.25	3.22	3.20	0.54
minority_religion	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.55
pray_private	7.91	8.14	8.18	7.97	7.97	7.97	7.81	0.55
electric_light	0.24	0.12	0.13	0.16	0.10	0.13	0.13	0.55
cement_floor	0.63	0.55	0.48	0.56	0.47	0.51	0.51	0.56
kerosene_light	0.24	0.33	0.31	0.26	0.31	0.29	0.37	0.56
travel_big_city	0.76	0.73	0.75	0.71	0.76	0.81	0.76	0.57
fumbira_lang	0.02	0.01	0.01	0.07	0.05	0.03	0.01	0.57
stone_wall	0.04	0.03	0.03	0.05	0.03	0.06	0.04	0.59
married	0.33	0.35	0.33	0.33	0.39	0.30	0.26	0.60
household_other	0.15	0.13	0.12	0.09	0.09	0.10	0.12	0.61
mufumbira_tribe	0.01	0.02	0.02	0.08	0.03	0.04	0.01	0.61
living_as_married	0.38	0.33	0.41	0.38	0.37	0.35	0.44	0.63
radio	0.81	0.85	0.79	0.84	0.85	0.82	0.83	0.64
number_children	3.25	3.05	3.27	3.13	3.77	3.51	3.32	0.67
single_hut	0.60	0.59	0.61	0.62	0.69	0.62	0.61	0.69
brick_wall	0.63	0.64	0.58	0.62	0.55	0.53	0.55	0.69
illiterate	0.10	0.07	0.11	0.09	0.11	0.12	0.08	0.69
christian_only	0.02	0.03	0.03	0.03	0.02	0.03	0.01	0.71
sofa	0.22	0.16	0.15	0.22	0.16	0.20	0.21	0.72
hospitality_work	0.06	0.07	0.03	0.04	0.04	0.06	0.04	0.72
several_huts	0.12	0.10	0.08	0.10	0.07	0.11	0.09	0.73
earth_floor	0.28	0.35	0.32	0.31	0.37	0.37	0.31	0.74
mud_wall	0.23	0.23	0.23	0.22	0.31	0.24	0.27	0.77
cement_wall	0.07	0.06	0.08	0.09	0.08	0.11	0.08	0.80
charcoal_fuel	0.46	0.46	0.36	0.36	0.36	0.35	0.40	0.80
firewood_fuel	0.51	0.51	0.60	0.60	0.62	0.62	0.58	0.80

share_house	0.27	0.31	0.31	0.27	0.24	0.27	0.31	0.82
retail_work	0.12	0.09	0.12	0.09	0.09	0.08	0.07	0.82
no_work	0.04	0.06	0.03	0.04	0.03	0.05	0.04	0.83
mutooro	0.01	0.02	0.01	0.00	0.02	0.01	0.02	0.84
agriculture_work	0.53	0.54	0.58	0.62	0.63	0.57	0.63	0.84
household_children	2.28	2.21	2.04	2.32	2.37	2.17	2.16	0.85
religious_service	1.89	2.02	1.55	1.14	1.79	1.34	1.52	0.86
household_spouse	0.21	0.23	0.27	0.26	0.26	0.25	0.22	0.86
living_conditions	0.04	0.01	0.04	0.08	0.04	0.03	-0.06	0.86
living_conditions_compared	1.95	2.03	1.96	2.04	2.06	2.01	1.95	0.88
members	4.44	4.31	4.17	4.40	4.51	4.19	4.23	0.88
household_younger	2.96	2.88	2.71	2.94	3.07	2.80	2.75	0.88
muslim	0.21	0.18	0.16	0.17	0.17	0.12	0.19	0.89
household_older	0.47	0.43	0.46	0.44	0.44	0.39	0.48	0.91
misc_fuel	0.04	0.03	0.03	0.04	0.02	0.03	0.03	0.92
muganda_tribe	0.55	0.58	0.54	0.55	0.51	0.52	0.60	0.92
day	1.25	1.19	1.21	1.22	1.25	1.24	1.25	0.96
household_head	0.64	0.64	0.62	0.64	0.65	0.65	0.66	0.99

Table 11: Balance on covariates among Compliers

	PLA	VAW	ABO	ABS	ABO_ABS	VAW_ABS	VAW_ABO	p-value
chair	0.85	0.85	0.85	0.90	0.87	0.85	0.79	0.01
living_conditions_compared	2.20	2.21	2.12	2.21	2.15	2.10	2.07	0.04
education_work	0.03	0.05	0.03	0.05	0.02	0.02	0.03	0.04
transport_work	0.04	0.04	0.03	0.03	0.02	0.05	0.01	0.07
stone_wall	0.03	0.01	0.04	0.03	0.03	0.06	0.05	0.09
living_conditions	0.13	0.10	0.15	0.19	0.09	0.08	0.07	0.11
catholic	0.40	0.36	0.35	0.45	0.42	0.44	0.36	0.12
cellphone	0.83	0.82	0.78	0.82	0.78	0.82	0.75	0.13
cement_wall	0.12	0.10	0.10	0.13	0.08	0.14	0.12	0.17
minority_lang	0.06	0.03	0.09	0.03	0.05	0.03	0.04	0.17
muslim	0.20	0.22	0.20	0.14	0.16	0.11	0.16	0.18
fumbira_lang	0.02	0.02	0.01	0.06	0.04	0.02	0.00	0.24
religious_service	2.08	1.95	1.97	1.27	1.57	1.50	1.78	0.27
highest_grade	6.95	7.25	7.15	7.07	6.54	6.70	6.74	0.28
minority_religion	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.29
female	0.47	0.49	0.44	0.47	0.51	0.47	0.48	0.29
atheist	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30
household_head	0.58	0.56	0.60	0.59	0.55	0.61	0.58	0.32
brick_wall	0.59	0.65	0.59	0.60	0.61	0.53	0.56	0.34
age	32.32	31.75	31.54	32.03	32.41	32.60	31.56	0.34
mukiga	0.04	0.02	0.04	0.01	0.04	0.04	0.02	0.35
runyannkole_lang	0.03	0.02	0.06	0.06	0.05	0.11	0.05	0.36
mufumbira_tribe	0.01	0.02	0.01	0.05	0.04	0.03	0.01	0.36
pray_private	8.10	8.18	8.05	8.17	8.15	8.14	7.94	0.37
household_spouse	0.33	0.35	0.34	0.33	0.38	0.32	0.35	0.37
household_children	2.73	2.79	2.55	2.64	2.77	2.50	2.80	0.38
day	1.30	1.20	1.22	1.23	1.24	1.25	1.19	0.40
luganda_lang	0.90	0.93	0.84	0.85	0.86	0.84	0.91	0.40
members	4.90	5.02	4.68	4.85	4.93	4.64	4.99	0.41
minority_tribe	0.17	0.16	0.12	0.15	0.12	0.10	0.14	0.41
other_work	0.06	0.06	0.04	0.05	0.07	0.04	0.05	0.44
single_hut	0.63	0.58	0.58	0.60	0.62	0.56	0.64	0.45
radio	0.82	0.81	0.85	0.82	0.83	0.79	0.79	0.45
mobile_phone_use	3.36	3.34	3.28	3.35	3.27	3.36	3.13	0.45

travel_big_city	0.72	0.68	0.73	0.68	0.72	0.74	0.69	0.46
household_older	0.49	0.55	0.46	0.48	0.53	0.46	0.49	0.49
english_christian	0.05	0.07	0.08	0.07	0.08	0.06	0.09	0.50
tv	0.28	0.30	0.23	0.27	0.24	0.24	0.21	0.52
separated	0.10	0.11	0.09	0.11	0.13	0.14	0.13	0.53
misc_floor	0.09	0.10	0.12	0.11	0.12	0.10	0.14	0.54
cement_floor	0.64	0.61	0.58	0.59	0.50	0.58	0.54	0.54
mutooro	0.01	0.01	0.02	0.00	0.01	0.01	0.02	0.54
misc_fuel	0.01	0.02	0.03	0.03	0.03	0.03	0.02	0.55
household_younger	3.41	3.47	3.20	3.37	3.40	3.18	3.50	0.55
read_only	0.03	0.03	0.02	0.03	0.04	0.03	0.03	0.55
munyoro	0.07	0.06	0.06	0.05	0.04	0.03	0.04	0.58
manual_work	0.09	0.06	0.08	0.10	0.09	0.06	0.07	0.59
share_house	0.27	0.32	0.31	0.29	0.26	0.31	0.26	0.61
rooms	2.63	2.74	2.70	2.87	2.74	2.64	2.70	0.61
write_and_read	0.82	0.84	0.81	0.81	0.78	0.81	0.81	0.61
munyarwanda	0.08	0.07	0.07	0.10	0.11	0.09	0.06	0.64
domestic_work	0.03	0.05	0.04	0.06	0.06	0.05	0.05	0.65
household_other	0.08	0.09	0.05	0.07	0.07	0.07	0.07	0.67
charcoal_fuel	0.43	0.42	0.43	0.37	0.36	0.35	0.36	0.68
married	0.38	0.43	0.39	0.37	0.40	0.37	0.35	0.68
university	0.04	0.05	0.06	0.05	0.04	0.04	0.02	0.68
electric_light	0.24	0.24	0.18	0.23	0.18	0.17	0.15	0.69
munyankole	0.08	0.09	0.12	0.10	0.09	0.16	0.11	0.69
earth_floor	0.27	0.29	0.30	0.31	0.38	0.32	0.31	0.70
not_married	0.14	0.15	0.15	0.16	0.12	0.16	0.13	0.72
living_as_married	0.40	0.34	0.40	0.38	0.36	0.37	0.42	0.75
same_village	0.42	0.40	0.37	0.42	0.37	0.38	0.43	0.75
several_huts	0.10	0.10	0.10	0.12	0.11	0.13	0.10	0.76
other_person	0.11	0.11	0.11	0.12	0.10	0.09	0.10	0.76
retail_work	0.14	0.14	0.13	0.10	0.14	0.12	0.13	0.76
misc_wall	0.02	0.02	0.05	0.02	0.03	0.03	0.04	0.77
misc_light	0.12	0.11	0.14	0.12	0.11	0.11	0.13	0.77
firewood_fuel	0.55	0.57	0.54	0.60	0.61	0.62	0.63	0.77
christian_only	0.02	0.03	0.03	0.02	0.03	0.02	0.03	0.79
protestant	0.20	0.20	0.20	0.18	0.19	0.22	0.23	0.80
solar_light	0.28	0.26	0.32	0.30	0.32	0.35	0.30	0.81
illiterate	0.11	0.09	0.12	0.11	0.12	0.12	0.10	0.81
number_children	4.39	4.24	4.31	4.28	4.57	4.32	4.49	0.84
sofa	0.27	0.28	0.24	0.26	0.24	0.22	0.25	0.86
hospitality_work	0.05	0.04	0.06	0.06	0.05	0.06	0.05	0.87
write_only	0.04	0.04	0.05	0.05	0.05	0.04	0.05	0.89
muganda_tribe	0.53	0.56	0.56	0.53	0.54	0.54	0.60	0.89
agriculture_work	0.54	0.55	0.59	0.55	0.55	0.60	0.60	0.94
motor_cycle	0.28	0.29	0.26	0.29	0.30	0.27	0.29	0.95
no_work	0.05	0.03	0.05	0.05	0.05	0.05	0.04	0.95
mud_wall	0.24	0.22	0.22	0.22	0.24	0.25	0.24	0.97
kerosene_light	0.26	0.28	0.26	0.27	0.29	0.26	0.29	0.98
holy_spirit	0.12	0.13	0.13	0.14	0.13	0.14	0.12	0.99

Table 12: Balance on covariates among Indirect Compliers

	PLA	VAW	ABO	ABS	ABO_ABS	VAW_ABS	VAW_ABO	p-value
married	0.47	0.44	0.32	0.37	0.51	0.35	0.49	0.00
separated	0.09	0.11	0.10	0.21	0.08	0.15	0.10	0.01
household_other	0.04	0.09	0.06	0.09	0.04	0.10	0.02	0.03

read_only	0.02	0.02	0.06	0.04	0.04	0.01	0.01	0.03
munyarwanda	0.08	0.09	0.06	0.10	0.09	0.07	0.16	0.03
household_older	0.51	0.71	0.57	0.51	0.58	0.59	0.54	0.04
education_work	0.07	0.04	0.02	0.02	0.01	0.05	0.05	0.04
highest_grade	6.90	7.72	6.80	6.50	5.96	7.48	6.51	0.05
number_children	4.71	4.15	4.53	4.39	5.30	4.10	4.90	0.07
university	0.08	0.07	0.04	0.04	0.02	0.09	0.06	0.10
mutooro	0.02	0.01	0.06	0.00	0.00	0.01	0.02	0.10
living_conditions_compared	2.26	2.37	2.15	2.31	2.19	2.23	2.35	0.11
luganda_lang	0.94	0.94	0.84	0.89	0.83	0.82	0.91	0.14
household_spouse	0.45	0.48	0.50	0.40	0.50	0.46	0.54	0.15
living_as_married	0.35	0.37	0.48	0.33	0.32	0.39	0.37	0.15
not_married	0.10	0.12	0.11	0.13	0.09	0.14	0.05	0.16
chair	0.83	0.88	0.77	0.83	0.83	0.86	0.84	0.17
retail_work	0.08	0.18	0.14	0.15	0.13	0.18	0.11	0.17
living_conditions	0.14	0.15	0.00	0.20	0.05	0.06	0.05	0.17
motor_cycle	0.20	0.31	0.30	0.28	0.26	0.25	0.30	0.19
share_house	0.29	0.38	0.34	0.28	0.24	0.40	0.31	0.20
misc_wall	0.01	0.01	0.04	0.00	0.05	0.01	0.04	0.20
age	33.17	32.81	31.91	33.68	33.65	32.30	34.37	0.23
munyoro	0.07	0.03	0.06	0.07	0.03	0.01	0.03	0.23
other_person	0.15	0.09	0.11	0.11	0.10	0.06	0.11	0.24
radio	0.77	0.85	0.75	0.82	0.78	0.80	0.82	0.25
household_head	0.51	0.43	0.43	0.52	0.45	0.44	0.44	0.25
minority_lang	0.04	0.04	0.10	0.02	0.06	0.04	0.03	0.25
single_hut	0.57	0.52	0.58	0.60	0.67	0.53	0.58	0.27
misc_floor	0.08	0.09	0.14	0.09	0.07	0.07	0.11	0.27
fumbira_lang	0.01	0.01	0.01	0.03	0.04	0.04	0.02	0.27
agriculture_work	0.55	0.46	0.60	0.57	0.52	0.44	0.57	0.29
mukiga	0.03	0.03	0.05	0.02	0.06	0.04	0.02	0.30
rooms	2.62	2.76	2.59	2.92	2.89	2.54	2.82	0.32
transport_work	0.02	0.03	0.02	0.01	0.04	0.01	0.03	0.32
earth_floor	0.25	0.20	0.28	0.29	0.36	0.33	0.26	0.34
other_work	0.06	0.07	0.02	0.05	0.08	0.04	0.06	0.34
mufumbira_tribe	0.03	0.01	0.00	0.02	0.03	0.04	0.01	0.36
muslim	0.18	0.23	0.17	0.12	0.22	0.13	0.18	0.41
munyankole	0.04	0.05	0.08	0.10	0.12	0.12	0.09	0.42
illiterate	0.18	0.10	0.14	0.11	0.16	0.13	0.18	0.45
runyannkole_lang	0.01	0.01	0.05	0.07	0.06	0.10	0.04	0.48
several_huts	0.14	0.09	0.08	0.11	0.08	0.08	0.12	0.49
domestic_work	0.11	0.08	0.08	0.09	0.07	0.11	0.10	0.51
minority_religion	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.55
travel_big_city	0.58	0.64	0.60	0.57	0.57	0.66	0.66	0.56
hospitality_work	0.04	0.06	0.04	0.07	0.06	0.08	0.03	0.57
mobile_phone_use	3.18	3.27	3.12	3.09	3.10	3.39	3.13	0.57
english_christian	0.09	0.04	0.06	0.09	0.08	0.06	0.08	0.59
kerosene_light	0.24	0.23	0.31	0.28	0.28	0.21	0.28	0.59
write_and_read	0.74	0.83	0.74	0.81	0.76	0.79	0.78	0.60
cement_wall	0.15	0.17	0.10	0.17	0.11	0.14	0.13	0.62
cement_floor	0.67	0.71	0.58	0.62	0.56	0.60	0.63	0.63
write_only	0.05	0.04	0.07	0.03	0.04	0.07	0.04	0.64
atheist	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.66
household_children	2.89	2.65	2.75	2.82	3.06	2.62	2.74	0.66
no_work	0.02	0.06	0.03	0.04	0.04	0.06	0.02	0.66
day	1.21	1.16	1.24	1.27	1.20	1.28	1.19	0.68
stone_wall	0.03	0.05	0.05	0.04	0.04	0.05	0.02	0.73
manual_work	0.07	0.06	0.07	0.04	0.06	0.09	0.06	0.73

same_village	0.31	0.30	0.28	0.36	0.28	0.31	0.31	0.74
protestant	0.17	0.17	0.21	0.17	0.17	0.20	0.19	0.75
minority_tribe	0.18	0.18	0.13	0.13	0.18	0.15	0.13	0.75
cellphone	0.78	0.81	0.82	0.76	0.77	0.80	0.77	0.76
household_younger	3.50	3.33	3.29	3.45	3.61	3.14	3.40	0.77
members	5.00	5.04	4.87	4.97	5.20	4.72	4.94	0.78
muganda_tribe	0.55	0.60	0.56	0.55	0.48	0.55	0.55	0.81
tv	0.31	0.37	0.28	0.31	0.28	0.31	0.25	0.82
catholic	0.37	0.36	0.36	0.45	0.36	0.40	0.35	0.84
female	0.61	0.64	0.66	0.60	0.63	0.64	0.65	0.84
mud_wall	0.20	0.13	0.24	0.16	0.21	0.19	0.18	0.86
firewood_fuel	0.54	0.55	0.55	0.60	0.59	0.56	0.63	0.86
electric_light	0.27	0.37	0.24	0.30	0.29	0.26	0.23	0.88
charcoal_fuel	0.45	0.45	0.44	0.39	0.40	0.42	0.37	0.88
holy_spirit	0.15	0.17	0.17	0.13	0.13	0.17	0.16	0.90
pray_private	8.24	8.21	8.19	8.20	8.00	8.09	8.20	0.90
misc_light	0.13	0.14	0.14	0.11	0.13	0.14	0.11	0.91
misc_fuel	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.91
religious_service	1.81	1.94	1.54	1.42	1.80	1.49	1.91	0.92
sofa	0.29	0.27	0.26	0.25	0.24	0.25	0.31	0.92
christian_only	0.03	0.02	0.02	0.03	0.02	0.02	0.03	0.93
solar_light	0.27	0.21	0.26	0.23	0.25	0.27	0.28	0.97
brick_wall	0.61	0.64	0.57	0.63	0.58	0.61	0.63	0.98

Table 13: Balance on covariates among Apprised Never-Takers

	PLA	VAW	ABO	ABS	ABO_ABS	VAW_ABS	VAW_ABO	p-value
motor_cycle	0.30	0.24	0.24	0.24	0.23	0.29	0.19	0.02
household_older	0.66	0.53	0.57	0.59	0.48	0.57	0.48	0.04
write_and_read	0.71	0.81	0.77	0.76	0.65	0.79	0.76	0.04
domestic_work	0.09	0.12	0.12	0.04	0.06	0.09	0.08	0.05
holy_spirit	0.18	0.14	0.16	0.17	0.16	0.22	0.12	0.06
not_married	0.09	0.12	0.15	0.12	0.15	0.18	0.10	0.06
cement_wall	0.16	0.13	0.16	0.16	0.14	0.24	0.15	0.07
travel_big_city	0.57	0.67	0.60	0.60	0.55	0.66	0.69	0.07
minority_lang	0.07	0.03	0.12	0.05	0.08	0.03	0.06	0.07
radio	0.79	0.83	0.83	0.77	0.70	0.79	0.79	0.09
female	0.58	0.53	0.60	0.60	0.51	0.53	0.51	0.13
living_conditions	0.09	0.13	0.11	0.09	-0.06	0.10	0.03	0.13
read_only	0.05	0.03	0.01	0.03	0.05	0.05	0.05	0.14
university	0.08	0.06	0.07	0.06	0.06	0.12	0.08	0.15
atheist	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.16
misc_light	0.15	0.17	0.17	0.16	0.20	0.12	0.17	0.19
write_only	0.06	0.04	0.05	0.05	0.07	0.02	0.05	0.19
rooms	2.73	2.64	2.64	2.92	2.56	2.80	2.78	0.22
cement_floor	0.61	0.73	0.66	0.63	0.52	0.72	0.60	0.23
pray_private	7.99	8.21	8.15	8.44	8.27	8.21	8.14	0.24
living_conditions_compared	2.18	2.26	2.27	2.24	2.05	2.27	2.14	0.24
stone_wall	0.02	0.04	0.07	0.04	0.03	0.04	0.03	0.25
brick_wall	0.53	0.64	0.51	0.58	0.53	0.52	0.59	0.25
highest_grade	6.99	7.22	6.93	6.76	5.95	7.69	6.97	0.25
illiterate	0.17	0.12	0.17	0.16	0.23	0.14	0.14	0.27
single_hut	0.62	0.52	0.64	0.59	0.68	0.58	0.62	0.29
luganda_lang	0.88	0.91	0.80	0.88	0.82	0.87	0.88	0.30
agriculture_work	0.50	0.50	0.56	0.63	0.60	0.47	0.58	0.30
household_children	2.88	2.66	2.42	2.92	2.71	2.62	2.70	0.33

catholic	0.36	0.31	0.34	0.37	0.42	0.33	0.33	0.34
household_head	0.49	0.56	0.48	0.52	0.59	0.54	0.54	0.34
runyannkole_lang	0.01	0.03	0.05	0.04	0.06	0.03	0.05	0.34
mukiga	0.03	0.03	0.04	0.01	0.06	0.02	0.03	0.35
manual_work	0.07	0.03	0.06	0.05	0.07	0.05	0.07	0.35
electric_light	0.23	0.27	0.26	0.26	0.17	0.41	0.23	0.36
mutooro	0.03	0.00	0.03	0.01	0.01	0.01	0.02	0.36
number_children	5.06	4.83	4.61	5.16	4.96	4.42	4.69	0.37
misc_floor	0.10	0.09	0.09	0.13	0.15	0.07	0.13	0.39
munyarwanda	0.10	0.11	0.09	0.08	0.11	0.13	0.05	0.39
household_spouse	0.43	0.34	0.43	0.40	0.34	0.36	0.40	0.43
several_huts	0.11	0.11	0.11	0.14	0.08	0.09	0.08	0.44
share_house	0.27	0.37	0.25	0.27	0.24	0.33	0.29	0.44
protestant	0.16	0.23	0.20	0.15	0.18	0.17	0.19	0.45
day	1.20	1.29	1.28	1.25	1.23	1.22	1.22	0.45
fumbira_lang	0.03	0.04	0.03	0.02	0.04	0.06	0.00	0.45
members	5.18	4.92	4.71	5.20	4.93	4.97	4.95	0.48
munyoro	0.07	0.03	0.07	0.07	0.07	0.02	0.05	0.48
earth_floor	0.29	0.18	0.26	0.24	0.33	0.21	0.26	0.51
minority_religion	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.54
other_person	0.09	0.06	0.08	0.11	0.10	0.07	0.08	0.54
sofa	0.36	0.30	0.25	0.28	0.24	0.34	0.31	0.55
household_other	0.09	0.11	0.09	0.08	0.07	0.10	0.06	0.55
kerosene_light	0.21	0.30	0.25	0.28	0.26	0.21	0.27	0.56
mufumbira_tribe	0.03	0.01	0.02	0.04	0.04	0.05	0.03	0.56
other_work	0.07	0.05	0.05	0.03	0.06	0.09	0.05	0.56
english_christian	0.05	0.08	0.09	0.08	0.07	0.06	0.10	0.57
chair	0.83	0.80	0.86	0.86	0.83	0.82	0.79	0.57
age	32.88	33.30	33.01	33.88	34.46	34.13	35.18	0.59
hospitality_work	0.04	0.07	0.03	0.04	0.04	0.04	0.03	0.59
tv	0.32	0.27	0.33	0.28	0.23	0.39	0.29	0.60
muganda_tribe	0.48	0.55	0.48	0.51	0.46	0.55	0.58	0.60
mobile_phone_use	3.20	3.39	3.27	3.08	3.08	3.45	3.17	0.60
charcoal_fuel	0.43	0.49	0.44	0.37	0.33	0.47	0.40	0.62
education_work	0.07	0.05	0.05	0.04	0.04	0.07	0.05	0.62
munyankole	0.07	0.08	0.11	0.08	0.10	0.06	0.07	0.66
household_younger	3.51	3.38	3.14	3.61	3.45	3.41	3.47	0.67
firewood_fuel	0.57	0.50	0.54	0.61	0.65	0.51	0.57	0.69
cellphone	0.77	0.82	0.83	0.75	0.75	0.85	0.79	0.70
transport_work	0.02	0.03	0.02	0.03	0.02	0.04	0.02	0.71
mud_wall	0.25	0.17	0.23	0.21	0.27	0.17	0.19	0.73
minority_tribe	0.20	0.21	0.17	0.20	0.16	0.17	0.18	0.75
muslim	0.20	0.19	0.15	0.18	0.13	0.17	0.18	0.76
solar_light	0.33	0.22	0.26	0.24	0.25	0.21	0.25	0.79
married	0.43	0.37	0.37	0.40	0.40	0.42	0.44	0.82
misc_fuel	0.00	0.01	0.02	0.02	0.02	0.02	0.03	0.85
separated	0.10	0.15	0.12	0.12	0.14	0.15	0.11	0.87
misc_wall	0.03	0.03	0.02	0.01	0.04	0.03	0.03	0.89
no_work	0.03	0.05	0.03	0.03	0.03	0.03	0.03	0.92
christian_only	0.04	0.06	0.06	0.04	0.04	0.04	0.07	0.94
religious_service	1.99	1.82	1.50	1.78	1.46	2.05	1.78	0.94
living_as_married	0.38	0.39	0.42	0.34	0.36	0.32	0.38	0.96
same_village	0.33	0.33	0.30	0.34	0.31	0.29	0.33	0.98
retail_work	0.13	0.15	0.11	0.12	0.12	0.13	0.11	0.99

Table 14: Balance on covariates among Never-Takers

	PLA	VAW	ABO	ABS	ABO_ABS	VAW_ABS	VAW_ABO	p-value
living_conditions	0.12	0.12	0.11	0.17	0.04	0.08	0.05	0.01
minority_lang	0.06	0.03	0.10	0.03	0.06	0.03	0.04	0.04
chair	0.84	0.84	0.83	0.87	0.85	0.84	0.80	0.06
minority_religion	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.08
not_married	0.12	0.14	0.14	0.14	0.12	0.16	0.10	0.08
cement_wall	0.14	0.12	0.12	0.15	0.10	0.16	0.13	0.12
highest_grade	6.95	7.35	7.02	6.85	6.26	7.14	6.74	0.13
age	32.68	32.32	32.02	32.94	33.22	32.94	33.20	0.15
write_and_read	0.77	0.83	0.78	0.79	0.75	0.80	0.79	0.16
muslim	0.19	0.21	0.18	0.15	0.16	0.13	0.17	0.18
transport_work	0.03	0.04	0.02	0.03	0.02	0.04	0.02	0.18
mutooro	0.02	0.01	0.03	0.01	0.01	0.01	0.02	0.19
education_work	0.05	0.05	0.03	0.04	0.02	0.04	0.04	0.20
living_conditions_compared	2.21	2.26	2.17	2.24	2.14	2.18	2.16	0.21
household_other	0.07	0.10	0.07	0.08	0.07	0.09	0.05	0.21
separated	0.10	0.12	0.10	0.14	0.12	0.14	0.12	0.21
single_hut	0.61	0.55	0.60	0.59	0.65	0.56	0.62	0.23
mukiga	0.03	0.02	0.04	0.01	0.05	0.03	0.02	0.23
household_spouse	0.39	0.38	0.40	0.37	0.40	0.36	0.41	0.25
luganda_lang	0.90	0.93	0.83	0.87	0.85	0.84	0.90	0.25
catholic	0.38	0.35	0.35	0.42	0.40	0.40	0.35	0.27
household_children	2.81	2.73	2.55	2.76	2.82	2.56	2.76	0.27
share_house	0.28	0.35	0.30	0.28	0.25	0.33	0.28	0.28
stone_wall	0.03	0.03	0.05	0.03	0.04	0.05	0.04	0.31
number_children	4.65	4.35	4.44	4.56	4.84	4.30	4.64	0.33
illiterate	0.15	0.10	0.14	0.13	0.16	0.13	0.13	0.33
travel_big_city	0.65	0.67	0.67	0.63	0.65	0.70	0.68	0.34
pray_private	8.11	8.20	8.11	8.25	8.14	8.15	8.05	0.35
runyannkole_lang	0.02	0.02	0.05	0.06	0.05	0.09	0.05	0.36
religious_service	1.99	1.92	1.75	1.45	1.59	1.64	1.81	0.38
members	5.00	5.00	4.73	4.98	4.99	4.75	4.97	0.38
household_younger	3.46	3.42	3.20	3.46	3.46	3.23	3.47	0.38
rooms	2.65	2.72	2.66	2.90	2.73	2.66	2.75	0.41
other_person	0.12	0.10	0.10	0.12	0.10	0.08	0.10	0.41
university	0.06	0.05	0.06	0.05	0.04	0.07	0.05	0.41
misc_fuel	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.42
married	0.41	0.42	0.37	0.38	0.43	0.38	0.41	0.43
english_christian	0.06	0.06	0.08	0.08	0.08	0.06	0.09	0.44
misc_floor	0.09	0.09	0.12	0.11	0.12	0.08	0.13	0.44
mobile_phone_use	3.27	3.33	3.25	3.21	3.18	3.39	3.14	0.44
cement_floor	0.64	0.66	0.60	0.61	0.52	0.62	0.58	0.45
household_older	0.54	0.58	0.52	0.52	0.53	0.52	0.50	0.45
other_work	0.06	0.06	0.04	0.04	0.07	0.05	0.05	0.46
cellphone	0.80	0.82	0.80	0.79	0.77	0.82	0.77	0.49
munyoro	0.07	0.05	0.06	0.06	0.05	0.03	0.04	0.49
brick_wall	0.58	0.64	0.57	0.60	0.58	0.54	0.58	0.51
living_as_married	0.38	0.36	0.42	0.36	0.35	0.36	0.40	0.52
earth_floor	0.27	0.25	0.28	0.28	0.36	0.29	0.29	0.53
minority_tribe	0.18	0.17	0.13	0.16	0.15	0.13	0.15	0.54
mufumbira_tribe	0.02	0.02	0.01	0.04	0.04	0.04	0.02	0.55
radio	0.80	0.82	0.82	0.81	0.79	0.79	0.80	0.56
manual_work	0.08	0.05	0.07	0.07	0.08	0.07	0.07	0.58
read_only	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.59
atheist	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62
munyankole	0.07	0.08	0.11	0.10	0.10	0.12	0.09	0.63

fumbira_lang	0.02	0.02	0.02	0.04	0.04	0.04	0.01	0.64
day	1.25	1.21	1.24	1.25	1.23	1.25	1.20	0.65
same_village	0.37	0.36	0.33	0.39	0.34	0.34	0.37	0.65
misc_light	0.13	0.13	0.15	0.13	0.14	0.12	0.14	0.73
charcoal_fuel	0.44	0.44	0.43	0.38	0.36	0.40	0.37	0.73
protestant	0.18	0.20	0.20	0.17	0.18	0.21	0.21	0.74
several_huts	0.11	0.10	0.10	0.12	0.10	0.11	0.10	0.77
tv	0.30	0.31	0.27	0.28	0.25	0.29	0.24	0.77
sofa	0.30	0.28	0.25	0.26	0.24	0.26	0.28	0.78
write_only	0.05	0.04	0.05	0.05	0.05	0.04	0.05	0.79
firewood_fuel	0.55	0.55	0.54	0.60	0.62	0.58	0.61	0.80
holy_spirit	0.15	0.14	0.15	0.14	0.14	0.17	0.13	0.82
female	0.54	0.53	0.53	0.54	0.54	0.53	0.53	0.82
retail_work	0.12	0.15	0.13	0.12	0.13	0.14	0.12	0.82
misc_wall	0.02	0.02	0.04	0.02	0.04	0.03	0.04	0.83
kerosene_light	0.24	0.27	0.27	0.27	0.28	0.24	0.28	0.83
muganda_tribe	0.52	0.57	0.54	0.53	0.50	0.55	0.58	0.83
christian_only	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.84
munyarwanda	0.09	0.09	0.07	0.09	0.10	0.10	0.08	0.84
electric_light	0.25	0.28	0.22	0.26	0.20	0.25	0.19	0.85
agriculture_work	0.53	0.52	0.58	0.58	0.56	0.53	0.59	0.85
household_head	0.54	0.53	0.53	0.55	0.53	0.55	0.53	0.90
hospitality_work	0.04	0.05	0.05	0.06	0.05	0.06	0.04	0.92
solar_light	0.29	0.24	0.29	0.27	0.29	0.30	0.28	0.94
mud_wall	0.23	0.19	0.23	0.20	0.24	0.21	0.21	0.95
domestic_work	0.07	0.07	0.07	0.06	0.06	0.07	0.07	0.96
no_work	0.04	0.04	0.04	0.04	0.04	0.05	0.03	0.99
motor_cycle	0.27	0.28	0.26	0.27	0.27	0.27	0.26	1.00

Table 15: Balance on covariates among Non-Compliers

D Compliance Strata by Treatment Status

We test whether the compliance status of respondents is affected by the treatment by computing a likelihood ratio permutation test. If a participant's response indicates they are a "Complier" we code their compliance variable 1, "Indirect Compliers" are coded as 2, "Apprised Never-Takers" as 3, and "Never-Takers" as 4. This variable is modeled as a multinomial logit process. We run a "full" model that specifies the self-selection into compliance strata as a function of treatment status, block fixed effects, and a resample indicator, and compare this to a "nested" model that restricts the coefficient on all treatment indicators to 0. We simulate 1000 log likelihood ratios between these models under the sharp null of no treatment effect on compliance choice for all units by permuting the treatment assignment. We then obtain p-values by calculating the proportion of simulated likelihood ratios that are at least as large as the observed likelihood ratio.

The results from this test are plotted on Figure 8. The observed likelihoods are not highly unlikely under the null hypothesis of no treatment effect on compliance choice for all units. We fail to reject the null at 5% confidence in all tests. Note that the p-value is below the $\alpha = .1$ threshold among women in the test that uses "pure" treatment categories.

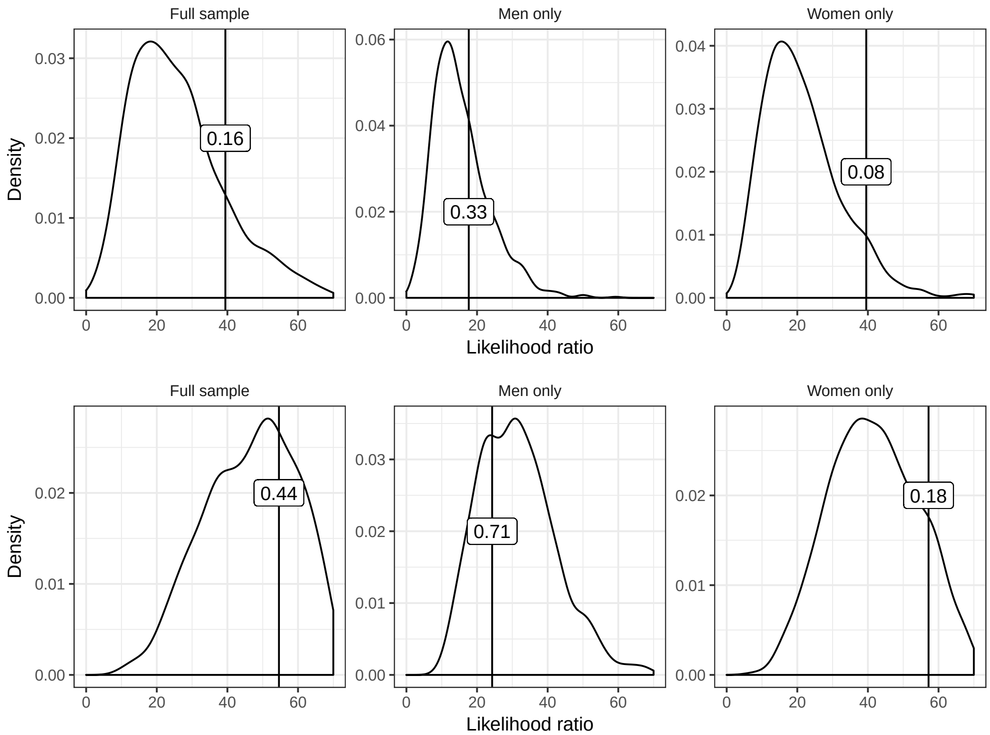

Figure 8: Tests of the assumption that treatment does not affect compliance.
Y axis plots the density of likelihood ratios simulated under the sharp null of no treatment effect on compliance status for all units. The X axis plots the value of the likelihood ratio. Vertical line indicates the observed likelihood ratio, and its label indicates the p-value. The first row derives from a full model in which treatment status is coded into "pure" treatment indicators (any VAW messaging, any abortion messaging, any absenteeism messaging). The second row derives from a full model in which treatment status is defined according to the seven conditions in the randomization: VAW, Absenteeism, Abortion, VAW + Absenteeism, VAW + Abortion, Abortion + Absenteeism, Placebo.

E Relaxing the Assumption of No Cross-Over Effects

Because our experiment does not feature an arm in which viewers were exposed to all three messages (VAW, absenteeism, and abortion stigma), the comparison of subjects that were assigned to a given message to subjects that were not assigned to that message is slightly imbalanced. Consider, for example, the VAW message. The VAW-untreated group is more likely to have been exposed to either absenteeism or abortion stigma treatments. In order to see how this imbalance arises, note that we have seven experimental conditions: placebo, VAW, absenteeism, abortion, VAW+absenteeism, VAW+abortion, and absenteeism+abortion. The VAW treatment group comprises VAW, VAW+absenteeism, and VAW+abortion, whereas the control group comprises the remaining four groups. The average marginal effect of the VAW message could be identified by comparing VAW to placebo, by comparing VAW+absenteeism to absenteeism, or by comparing VAW+abortion to abortion. Because we do not have an VAW+absenteeism+abortion group, we do not have a treated counterpart to the absenteeism+abortion control group.

An alternative estimator to the one used in this paper simply excludes the absenteeism+abortion group and includes a fixed effect for the VAW+absenteeism and absenteeism groups (to control for the effects of the absenteeism treatment) and a fixed effect for the VAW+abortion and abortion groups (to control for the effects of the abortion treatment). This alternative estimators for the absenteeism and abortion treatments are analogous. As can be seen in tables 16 to 19, we find that these estimators produce very similar point estimates.

	Dependent variable:				
	Index of willingness to take action to counter absenteeism				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
absenteeism	0.045^{***}	-0.002	0.016	-0.002	0.0003
	(0.014)	(0.009)	(0.014)	(0.011)	(0.007)
abortion	-0.009	0.024^{**}	0.008	-0.013	0.012
	(0.017)	(0.011)	(0.018)	(0.014)	(0.009)
VAW	0.020	0.017	-0.013	-0.034^{**}	-0.003
	(0.020)	(0.011)	(0.017)	(0.013)	(0.008)
Control Mean	0.61	0.59	0.59	0.6	0.59
Vill. Means	0.62	0.6	0.59	0.58	0.59
Vill. SD	0.1	0.08	0.11	0.11	0.07
N Vill.	142	142	142	142	142
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,244	2,817	1,279	1,463	5,559
Adjusted R^2	0.024	0.044	0.053	0.070	0.048
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 16: Direct effects and spillovers from absenteeism messages among all respondents in endline surveys following 2015 and 2016 festivals using an estimator that is unbiased in the presence of cross-over effects.

Coefficients estimated using least-squares regression, conditioning on block fixed-effects and an indicator for resampling. All analyses exclude respondents from clusters assigned to the VAW+abortion treatment condition. *VAW* is a fixed effect for the VAW+absenteeism and VAW groups and *abortion* is a fixed effect for the absenteeism+abortion and abortion groups. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:				
			Education is an important goal		
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
absenteeism	0.075***	0.013	−0.045	0.005	−0.001
	(0.025)	(0.017)	(0.024)	(0.024)	(0.012)
abortion	−0.035	−0.007	−0.006	0.039	0.002
	(0.029)	(0.020)	(0.031)	(0.028)	(0.014)
VAW	0.008	−0.047**	0.040	0.035	−0.007
	(0.034)	(0.021)	(0.030)	(0.032)	(0.014)
Control Mean	0.42	0.43	0.46	0.42	0.43
Vill. Means	0.41	0.43	0.46	0.43	0.43
Vill. SD	0.19	0.12	0.18	0.17	0.08
N Vill.	142	142	142	142	142
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,244	2,817	1,279	1,463	5,559
Adjusted R^2	0.007	0.0002	0.0001	0.009	−0.0001
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 17: Direct effects and spillovers from absenteeism messages among all respondents in endline surveys following 2015 and 2016 festivals using an estimator that is unbiased in the presence of cross-over effects.

Coefficients estimated using least-squares regression, conditioning on block fixed-effects and an indicator for resampling. All analyses exclude respondents from clusters assigned to the VAW+abortion treatment condition. *VAW* is a fixed effect for the VAW+absenteeism and VAW groups and *abortion* is a fixed effect for the absenteeism+abortion and abortion groups. Standard errors are clustered at the village level. Two-tailed *p*-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:				
	Index of willingness to take action to counter intimate partner violence				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
VAW	0.058^{***}	0.001	-0.027	0.009	-0.002
	(0.016)	(0.012)	(0.019)	(0.017)	(0.010)
absenteeism	-0.026	-0.016	0.013	0.012	-0.003
	(0.020)	(0.013)	(0.024)	(0.025)	(0.013)
abortion	-0.053^{**}	0.013	0.003	-0.013	0.006
	(0.021)	(0.014)	(0.024)	(0.024)	(0.012)
Control Mean	0.37	0.38	0.4	0.37	0.38
Vill. Means	0.37	0.37	0.38	0.37	0.38
Vill. SD	0.1	0.07	0.13	0.13	0.06
N Vill.	94	94	94	93	94
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	954	2,101	831	849	3,781
Adjusted R^2	0.027	0.007	0.021	0.020	0.011
Notes:					*p<0.1; **p<0.05; ***p<0.01

Table 18: Direct effects and spillovers from anti-VAW messages among all respondents in endline surveys following 2016 festival using an estimator that is unbiased in the presence of cross-over effects.

Coefficients estimated using least-squares regression, conditioning on block fixed-effects and an indicator for resampling. All analyses exclude respondents from clusters assigned to the absenteeism+abortion treatment condition. *absenteeism* is a fixed effect for the VAW+absenteeism and absenteeism groups and *abortion* is a fixed effect for the VAW+abortion and abortion groups. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

	Dependent variable:				
	Willingness to help someone suffering from post-abortion complications				
	Compliers	Indirect Compliers	Apprised Never-Takers	Never-Takers	All Non-Compliers
	(1)	(2)	(3)	(4)	(5)
abortion	0.057^{***}	-0.007	-0.031	0.020	-0.005
	(0.021)	(0.014)	(0.023)	(0.023)	(0.012)
absenteeism	-0.029	-0.007	0.024	-0.013	-0.004
	(0.022)	(0.018)	(0.026)	(0.028)	(0.014)
VAW	-0.042	0.009	0.058^{**}	0.035	0.023
	(0.024)	(0.017)	(0.027)	(0.028)	(0.014)
Control Mean	0.81	0.8	0.8	0.77	0.79
Vill. Means	0.78	0.8	0.8	0.78	0.79
Vill. SD	0.17	0.12	0.17	0.18	0.11
N Vill.	143	143	143	142	143
Block FE	Yes	Yes	Yes	Yes	Yes
Observations	1,294	2,842	1,265	1,465	5,572
Adjusted R^2	0.007	0.017	0.031	0.015	0.020
Notes:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 19: Spillovers from anti-abortion stigma messages among all respondents in endline surveys following 2015 and 2016 festivals using an estimator that is unbiased in the presence of cross-over effects.

Coefficients estimated using least-squares regression, conditioning on block fixed-effects and an indicator for resampling. All analyses exclude respondents from clusters assigned to the VAW+absenteeism treatment condition. *VAW* is a fixed effect for the VAW+abortion and VAW groups and *absenteeism* is a fixed effect for the absenteeism+abortion and absenteeism groups. Standard errors are clustered at the village level. Two-tailed p-values are calculated by comparing the observed estimate to 2000 estimates simulated under the sharp null of no effects for all units by permuting the treatment assignment 2000 times.

F An Alternative High-Powered Test

Given the null results on spillovers, we seek here to understand what sorts of positive spillover effects we are able to rule out when we conduct a joined test across all three issue domains. Such a test will be better powered than three tests of three separate null hypotheses, each pertaining to just one of the issue domains.

Continuing with the notation employed in equation 6, we consider models of the form

$$\mathbf{Y}^m = \gamma_0^m + \mathbf{B}\boldsymbol{\gamma}^m + \tau_s \mathbf{z}^m + \mathbf{X}\boldsymbol{\delta}^m + \boldsymbol{\epsilon}^m. \tag{6}$$

We seek to test hypotheses that stipulate the same constant treatment effect τ_s for all messaging treatments m and across all respondents in stratum s. In effect, we posit that education-entertainment exerts a true underlying average treatment effect across subjects and issue domains.

Differencing out block and resample fixed effects gives us the residualized outcomes,

$$\hat{\mathbf{Y}}^m = \mathbf{Y}^m - (\hat{\gamma}_0^m + \mathbf{B}\hat{\boldsymbol{\gamma}}^m + \mathbf{X}\hat{\boldsymbol{\delta}}^m) \qquad = \tau_s \mathbf{z}^m + \boldsymbol{\epsilon}^m.$$

Suppose we hypothesise a constant marginal effect for participants in stratum s, denoted τ_s^h. If the data were truly generated by τ_s^h, it follows that

$$\hat{\mathbf{Y}}^m - \tau_s^h \mathbf{z}^m = \tilde{\mathbf{Y}}_{z=0}^m = \boldsymbol{\epsilon}.$$

In other words, if the data was generated by a stratum-specific constant effect, we can subtract that effect from the observed outcome of those in treatment and thereby obtain the vector of outcomes we would have observed if all units were put into control (Rosenbaum, 2002). For a given stratum s, we denote this 'hypothesized' control outcome implied under hypothesis τ^h by $\hat{\mathbf{Y}}_{z=0|h}^m$.

When hypothesis τ_s^h is true, the sharp null hypothesis of no treatment effect of $\mathbf{z}^m$ on $\hat{\mathbf{Y}}_{z=0|h}^m$ for all units will also be true because subtracting $\tau_s^h \mathbf{z}^m$ from $\hat{\mathbf{Y}}^m$ removes any systematic relationship between the treatment and the hypothesized control outcome. This implies we can test for the consistency of the data with τ_s^h by testing the null hypothesis of no effects of $\mathbf{z}_m$ on $\hat{\mathbf{Y}}_{z=0|h}^m$.

We employ the sum of squared residuals (SSR) as a test statistic (Bowers, Fredrickson, and Aronow, 2016). Denoting by s' the stratum of interest, $\hat{\tau}_s$ the estimate of the treatment effect arising from an OLS regression of $\hat{\mathbf{Y}}_{z=0|h}^m$ on $\mathbf{z}^m$ among stratum s', and α the intercept from this regression, we define the observed SSR test statistic as

$$\mathcal{T}(\hat{\mathbf{Y}}_{z=0|h}^m, \mathbf{z}^m) = \sum_m \sum_{i:s=s'} (\hat{\mathbf{Y}}_{z=0|h}^m - \mathbf{z}^m \hat{\tau}_s - \alpha)^2.$$

Note that here, we sum the squared residuals across the three regressions of outcome m on treatment m.

We can compare this observed test statistic to the distribution of the test statistic under the sharp null of no effect of $\mathbf{z}^m$ on $\hat{\mathbf{Y}}_{z=0|h}^m$ by generating 1000 permutations of the treatment assignment vector. Denoting one randomly permuted assignment vector by $\mathbf{z}'^m$, our p-value for the null hypothesis of no effect of $\mathbf{z}^m$ on $\hat{\mathbf{Y}}^m_{z=0|h}$ can be defined as:

$$E[\mathcal{T}(\hat{\mathbf{Y}}^m_{z=0|h}, \mathbf{z}^m) \geq \mathcal{T}(\hat{\mathbf{Y}}^m_{z=0|h}, \mathbf{z}'^m)].$$

As the data become less (more) plausible under τ^h_s, this p-value will approach 0 (1).

We apply this testing procedure to the endline data using the outcomes presented in Tables 3, 5 and 6. We first residualize the outcomes for block and resample effects, before standardizing them to have a standard deviation of one. As specified in our pre-registration plan, we consider only positive hypotheses. Specifically, we test hypotheses about constant stratum-specific effects ranging from 0 to .5 standard deviations. To maximize the power of the test we consider effects only among two compliance strata: Compliers and "Non-Compliers" (Indirect Compliers, Apprised Never-Takers and Never Takers).

Before delving into the results it is helpful to know the point estimates of each treatment on each standardized outcome. We estimate a .13 standard deviation (SD) effect of the absenteeism messaging on the absenteeism conative attitude index for compliers and a -.01 effect for non-compliers, a .08 SD effect of the abortion messaging on willingness to get involved for compliers and a -.05 effect for non-compliers, and a .18 SD effect of the VAW messaging on willingness to report incidents of VAW for compliers, with a .01 effect for non-compliers. For the purpose of this exercise, we think of these effects as arising with sampling variability from the same underlying constant stratum-specific effect.

The plot below presents results for Compliers, with hypothesized treatment effects on the horizontal and the corresponding p-values on the vertical axis. The dashed lines corresponds to the $\alpha = .10$ and $\alpha = .05$ confidence levels. Thus, any hypothesis whose p-value falls below these lines can be rejected with 90% and 95% confidence, respectively.

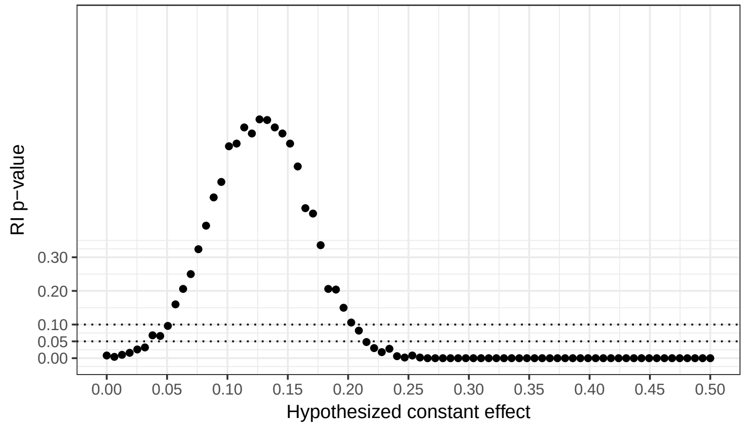

Figure 9: Constant Effect Hypotheses among Compliers

The analysis suggests that our data are most consistent with the hypothesis of a roughly .13 SD constant effect among compliers across the three messaging strategies. With 90% confidence, we can reject the hypothesis of a constant effect any smaller than .05 SD and any larger than .20 SD. With 95% confidence we can reject hypotheses of a constant effect smaller than .03 SD and greater than .22.

Turning now to Non-Compliers, we are able to rule out a range of constant effect hypotheses. Specifically, with 95% confidence we can rule out hypotheses constant spillover effects larger than .03 SD. The analysis thus suggests that spillovers, if they do occur, are too small to be of much substantive interest.

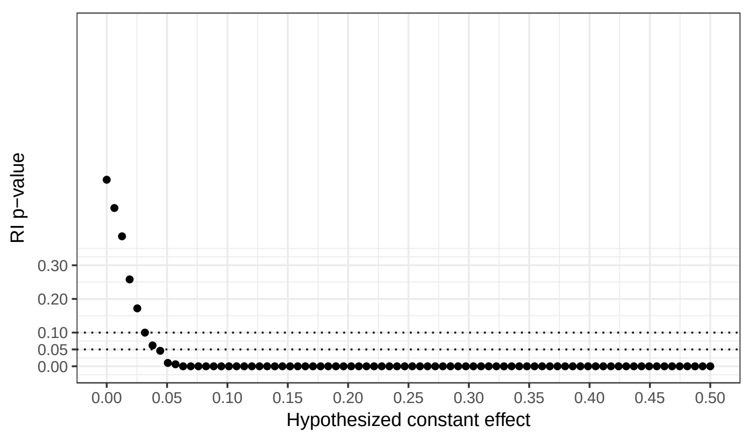

Figure 10: Constant Effect Hypotheses among Non-Compliers