

Addendum to Pre-Analysis Plan for ‘The Effects of Media Messages on Social Attitudes in Uganda’

Susanne Baltes, Jasper Cooper, Donald Green, Anna Wilke

March 4, 2016

The purpose of this addendum is to pre-specify the analysis and collection of new data for the project ‘The Effects of Media Messages on Social Attitudes in Uganda’. Upon first analysis of the endline data, collected in February 2016, we observed that some clusters in our survey contained only a very small number of direct compliers (those who attended at least one of the screenings).

1 Sampling strategy

In order to increase the number of compliers, we are surveying 15 additional respondents in 14 clusters. To select the 14 clusters, we identified the two clusters in each of the 7 treatment conditions with the fewest direct compliers. In one such case, the three clusters with the lowest number of compliers had the same number of compliers. We randomly selected two among the three. In another case, two clusters had numbers of compliers equal to the second lowest number of compliers for that condition. We randomly selected one of these, in order to obtain the two clusters with the lowest number of compliers. We have included the exact code used to perform this selection below. Note that this code was run on all data that had been collected as of February 2016, which did not include a small number of respondents that had yet to be tracked by the enumerators.

```
library(dplyr)

# Get number of compliers per TC -----

set.seed(1:7)

compliers <-
  el %>%
  group_by(tc) %>%
  summarize(
    direct = sum(compliance == 2, na.rm = T),
    treatment = treatment[1]
  ) %>%
  ungroup() %>%
  arrange(treatment, direct) %>%
  group_by(treatment) %>%
  summarize(
    tc_1 = tc[1],
    direct_1 = direct[1],
    tc_2 = tc[2],
    direct_2 = direct[2],
    tc_3 = tc[3],
    direct_3 = direct[3],
    tc_4 = tc[4],
    direct_4 = direct[4]
  ) %>%
```

```

arrange(treatment) %>%
ungroup() %>%
mutate(
  all_tie =
    ((direct_1 == direct_2) &
     (direct_2 == direct_3) &
     (direct_1 == direct_3)),
  top_tie =
    ((direct_2 == direct_3) &
     direct_1 != direct_2)
)

chosen <-
compliers %>%
filter(!all_tie & !top_tie) %>%
select(treatment,
       tc_1, tc_2)

top_ties <-
compliers %>%
filter(top_tie) %>%
mutate(tc_2 = sample(x = c(as.character(tc_2),
                          as.character(tc_3)),
                    size = 1)) %>%
select(treatment,tc_1,tc_2)

all_ties <-
compliers %>%
filter(all_tie) %>%
select(treatment,tc_1,tc_2,tc_3)

for(all_tie in 1:nrow(all_ties)){
  all_ties[all_tie,c("tc_1","tc_2")] <-
    sample(x = all_ties[all_tie,c("tc_1","tc_2","tc_3")],
          size = 2)
}

resample <-
rbind(
  chosen,top_ties,all_ties[,1:3]
)

kable(resample)

```

In the main round of endline data collection, we listed households in the catchment area of the video hall and put them in a random order, surveying the first 40 from this list, using the subsequent households as replacements wherever the first 40 did not meet eligibility requirements (see main Pre-Analysis Plan for more details). In this second round of sampling for compliers, households are selected from the same randomly ordered list, continuing the sequence followed 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.

2 Updates to estimation strategy

The new data collected will be used in sub-group analyses conducted with our HLM estimator as described in the pre-analysis plan. We will add a stratum dummy variable that takes the value 1 for all respondents that have been sampled in the second round and include it in all HLM models.

The code below also corrects a small bug in the HLM subgroup modeling code in the original PAP. We originally specified subgroup analyses using the covariates from the SUR models. In subgroup analyses we instead use the correct individual-level covariates used in the other HLM specifications.

```
make_HLM_formula <- function(  
  outcome = NULL,  
  treatment_indicator = NULL,  
  baseline_index = NULL,  
  crossover_null_rejected = FALSE  
)  
{  
  el_covariates_col <- paste(el_covariates,collapse = " + ")  
  tc_level_covariates_col <- paste(tc_level_covariates,collapse = " + ")  
  if(crossover_null_rejected){  
    treatment_indicator <- "IPV + absenteeism + abortion"  
  }  
  formula_core <- paste0(  
    outcome," ~ 1 + (1|tc) + ",treatment_indicator  
  )  
  HLM_formula <- paste0(  
    formula_core," + block + ",  
    el_covariates_col," + ",  
    tc_level_covariates_col," + ",  
    baseline_index  
  )  
  return(as.formula(HLM_formula))  
}  
  
ABO_attitude_formula <- make_HLM_formula(  
  outcome = "ABO_attitude_index",  
  treatment_indicator = "abortion",  
  baseline_index = "ABO_attitude_index_bl",  
  crossover_null_rejected = FALSE  
)  
  
ABO_attitude_HLM_complier <- lmer(  
  formula = ABO_attitude_formula,  
  data = subset(el,compliance ==2)  
)
```