Replication code for 'Voter information campaigns and political accountability: Cumulative findings from a preregistered meta-analysis of coordinated trials'

February 22, 2019

Contents

1 Notes on replication script

- This .Rmd file produces all tables, and figures reported in Dunning, Thad et al. 2019. "Voter information campaigns and political accountability: Cummulative findings from a preregistered meta-analysis of coordinated trials." *Science Advances*.
- The code exports results to folders "tables" and "figures" in this file's directory.
- Replication options include:
  - `prep`: if `TRUE`, generate data from raw files, if `FALSE`, loads cleaned data
  - `n_iter`: number of permutations of treatment assignment used in randomization inference analysis.
  - `load_iter_matrix`: if `FALSE`, generate `n_iter` columns of treatment assignment according to each study's assignment strategy. If `TRUE`, load existing assignment matrix with 2,000 iterations of treatment assignment.
  - `ri_action`: treatment effects and p-values from randomization inference can be generated from scratch if set to `"generate"`, loaded from existing file if (`"load"`), or not returned (`"ignore"`) in results. Note that these estimations may take a few hours depending on the value of `n_iter`.
  - `run_bayes`: if `TRUE`, will run Monte Carlo simulations. `run_bayes <- FALSE` will skip the Bayesian analysis.
  - `load_specification_matrices <- FALSE` will estimate, from scratch, treatment effects under all combinations of specification choices. Likewise, `load_mde_data <- FALSE` will re-run the diagnostic process for the minimal detectable effect analysis. Setting these options to `FALSE` will take several hours to run on an average machine.
  - `full_m1`: if `FALSE` turns off imputing `0`s for vote choice for cases with where voter turnout is 0 in the cleaned data.
- For a full clean replication empty the folders `figures` and `tables` and subfolder `data/temp` and set options: `prep <- TRUE`, `ri_action <- "generate"`, `run_bayes <- TRUE`, `load_iter_matrix <- FALSE`, `load_specification_matrices <- FALSE`, `load_mde_data <- FALSE`.
- This code requires installation of Rstan (unless `run_bayes <- FALSE`). See installation instructions here.
- The package `xlsx` requires a 64 bit Java install on PCs; for more detail see https://java.com/en/download/win10.jsp.

```
prep <- FALSE
n_iter <- 2000
load_iter_matrix <- TRUE
ri_action <- "generate"
run_bayes <- FALSE
load_specification_matrices <- FALSE
load_mde_data <- TRUE
full_m1 <- TRUE
```

```
if(prep){
  madat <- pool_studies(
    add_vars= c("nij", "Q_alt", "N_alt", "lc5.councillor.party.switch",
                "lc5.chair.party.switch", "lc5.councillor.redistricted2016"))
  if(full_m1){
    madat$m1[is.na(madat$m1) & madat$m3==0] <- 0
    #m3==0 in Burkina Faso means unlikely to vote or definitely not voting.
    madat$m1[madat$m1==1 & madat$m3==0 & madat$ctry == "bf"] <- 0
    madat_raw <- pool_studies(dir = "data/temp")
    madat$m1_against <- ifelse(madat_raw$m1==0 & madat$m3==1, 1, 0)
    madat$m1_against[is.na(madat$m1_against) & madat$m3==0 & madat$ctry != "bf"] <- 0
    madat$m1_against[madat$m1==1 & madat$m3==0 & madat$ctry == "bf"] <- 0
    madat$m1NA <- (is.na(madat$m1) & madat$m3 == 1)

  }
  save(madat, file = "data/temp/madat.Rda")
  write.csv(madat, "madat.csv")
}

load("data/temp/madat.Rda")

# Load previously generated permutation matrices
if(load_iter_matrix){
  load("data/temp/assign_iter.Rda")
}else{
  system.time(assign_iter <- assign_i(madat, n_iter))
  save(assign_iter, file = "data/temp/assign_iter.Rda")
}
  madat_iter <- cbind.data.frame(madat, assign_iter)
  rm(assign_iter)
  save(madat_iter, file = "data/temp/madat_iter.Rda")
```

2 Figure 1: Prior Beliefs and Politician Performance

3 Figure 2: Treatment Effect of Information on Incumbent Vote Choice

```
  cov <- "m14i+m17i+m18i+m20i+m21i+m22i+m24i+m26i+m27i"

  t1c1 <-  results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                   file_rite_obj = "data/temp/ri_m1pool_all_g_unadj.Rda",
                   good = TRUE, depvar = "m1", exclude_councilors = FALSE)
  t1c2 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                   file_rite_obj = "data/temp/ri_m1pool_ben_g_unadj.Rda",
                   depvar = "m1", weights = FALSE, good = TRUE, country = "ben")
  t1c3 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                   file_rite_obj = "data/temp/ri_m1pool_brz_g_unadj.Rda",
                   depvar = "m1", weights = TRUE,  good = TRUE, country = "brz")
  t1c4 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                   file_rite_obj = "data/temp/ri_m1pool_bf_g_unadj.Rda",
```

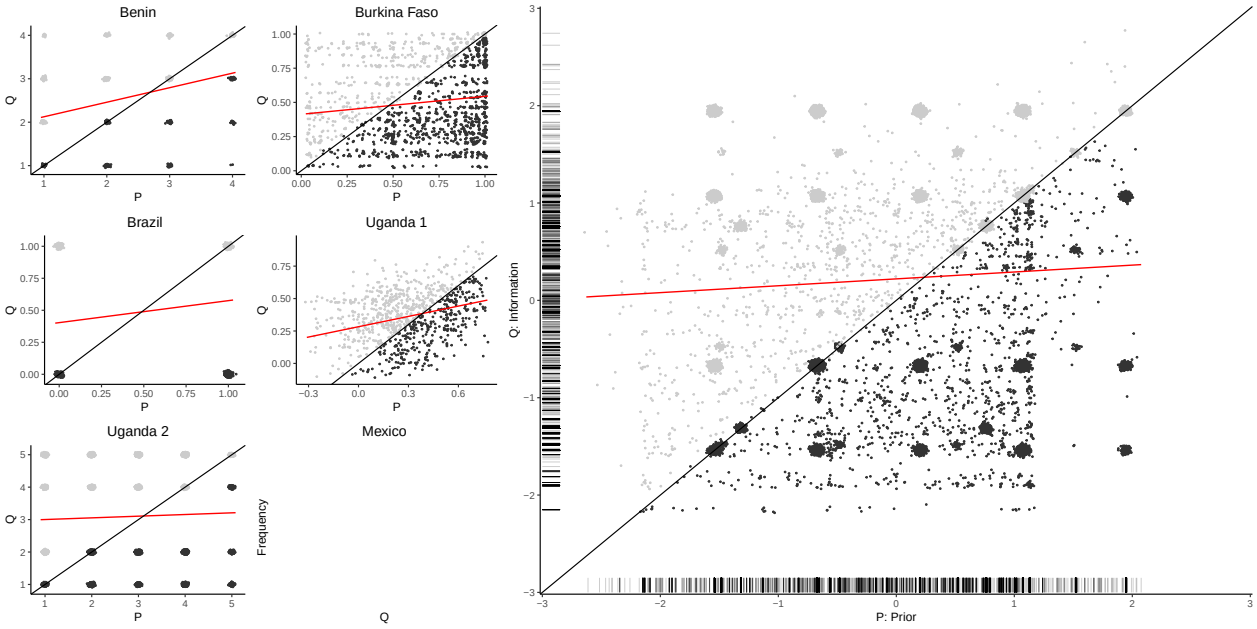

Figure 1: Prior Beliefs and Politician Performance. The figure plots performance information (Q) against prior beliefs (P) in each of the studies (left side) and across all studies (right side). Voters are in the good news group (gray) if information exceeds priors ($Q > P$), or if it confirms positive priors ($P = Q$, and Q is greater than median); otherwise, they are in the bad news group (black). On the right side, P and Q are standardized with a mean of 0 and standard deviation of 1 in each study. The density of the dotted areas is proportionate to the number of voters at each value of P and Q; for the pooled analysis, the rugs along the horizontal and vertical axes indicate the distribution of values. The Mexico study lacked a pre-intervention survey, thus we determine the good news and bad news groups according to whether Q is greater than median. The red lines indicate the linear fit between priors and information. For the pooled analysis, the slope of the fit is 0.071; the correlation is 0.053.

```
                depvar = "m1", weights = FALSE, good = TRUE, country = "bf")
t1c5 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_mex_g_unadj.Rda",
                depvar = "m1", weights = FALSE, good = TRUE, country = "mex")
t1c6 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_ug1_g_unadj.Rda",
                depvar = "m1", weights = FALSE, good = TRUE, country = "ug1")
t1c7 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_ug2_g_unadj.Rda",
                depvar = "m1", weights = TRUE,  good = TRUE, country = "ug2",
                exclude_councilors =FALSE)

 t2c1 <-  results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_all_b_unadj.Rda",
                good = FALSE, depvar = "m1", exclude_councilors =FALSE)
t2c2 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_ben_b_unadj.Rda",
                depvar = "m1", weights = FALSE, good = FALSE, country = "ben")
t2c3 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_brz_b_unadj.Rda",
                depvar = "m1", weights = TRUE,  good = FALSE, country = "brz")
t2c4 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_bf_b_unadj.Rda",
                depvar = "m1", weights = FALSE, good = FALSE, country = "bf")
t2c5 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_mex_b_unadj.Rda",
                depvar = "m1", weights = FALSE, good = FALSE, country = "mex")
t2c6 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_ug1_b_unadj.Rda",
                depvar = "m1", weights = FALSE, good = FALSE, country = "ug1")
t2c7 <- results(dat = madat_iter, ri_p = ri_action, sims = n_iter,
                file_rite_obj = "data/temp/ri_m1pool_ug2_b_unadj.Rda",
                depvar = "m1", weights = TRUE,  good = FALSE, country = "ug2",
                exclude_councilors =FALSE)
```

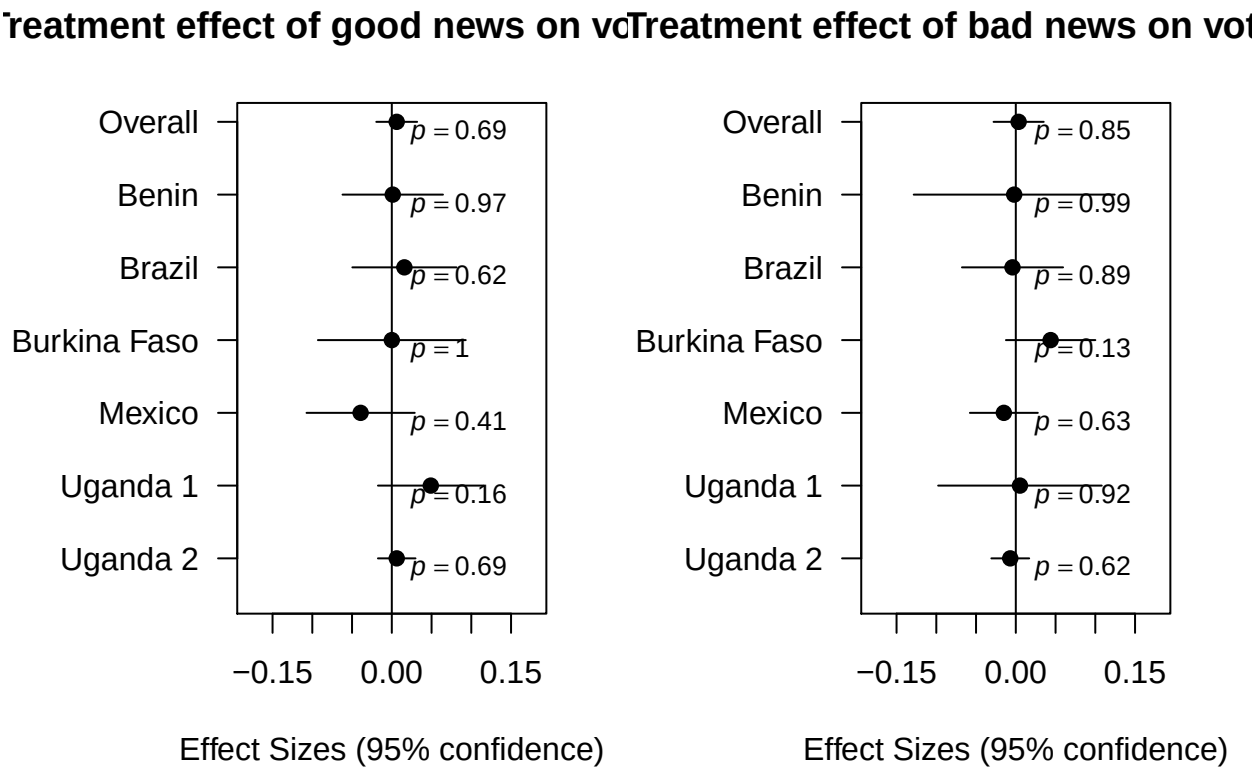

Figure 2: Meta-Analysis: Country-Specific Effects on Vote Choice. Estimated change in the proportion of voters who support an incumbent after receiving good news (left panel) or bad news (right panel) about the politician, compared to receiving no information. Unadjusted estimates. For estimating the average of the study-specific effects (top row), each study is weighted by the inverse of its size. Horizontal lines show 95% confidence intervals for the estimated change. Entries under each estimate show p-values calculated by randomization inference. In all cases, the differences are close to zero and statistically insignificant.

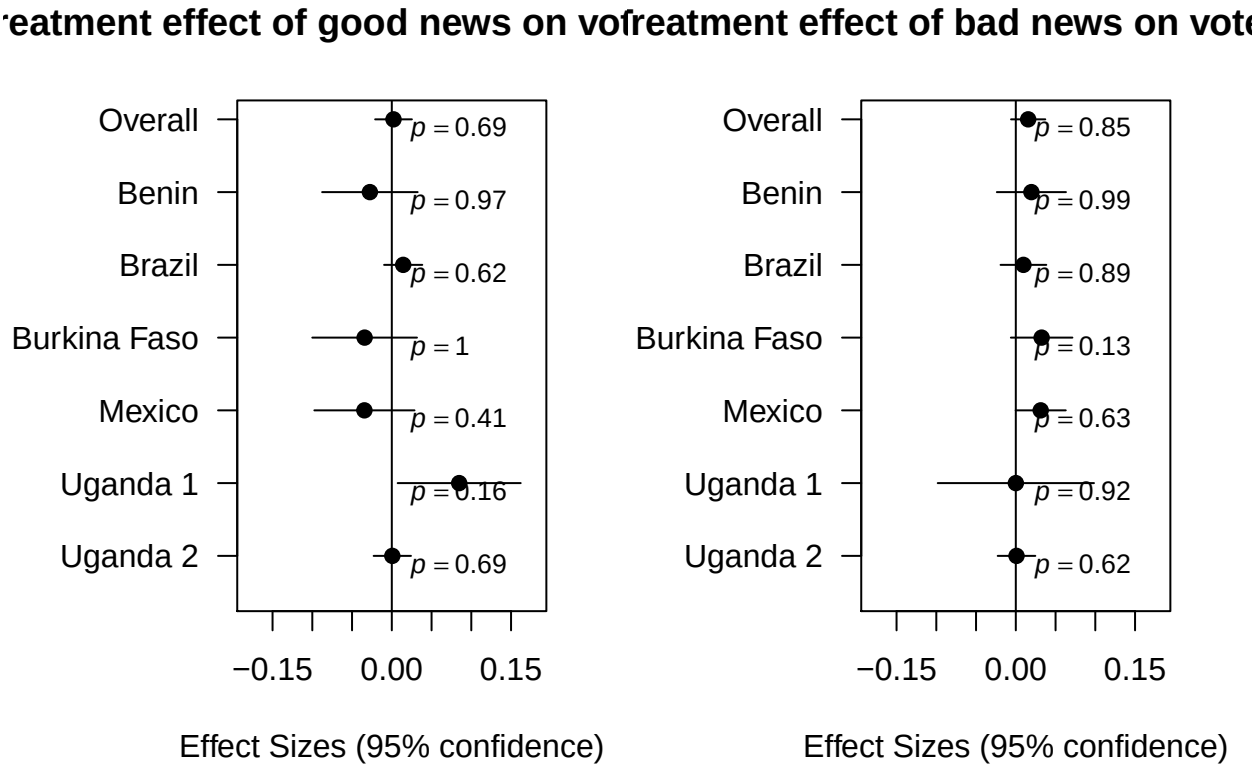

Figure 3: Meta-Analysis: Country-Specific Effects on Turnout. See notes to Figure **??**. In all but one test, the differences are close to zero and are statistically insignificant, using p-values from randomization inference.

4 Figure 3: Treatment Effect of Information on Voter Turnout

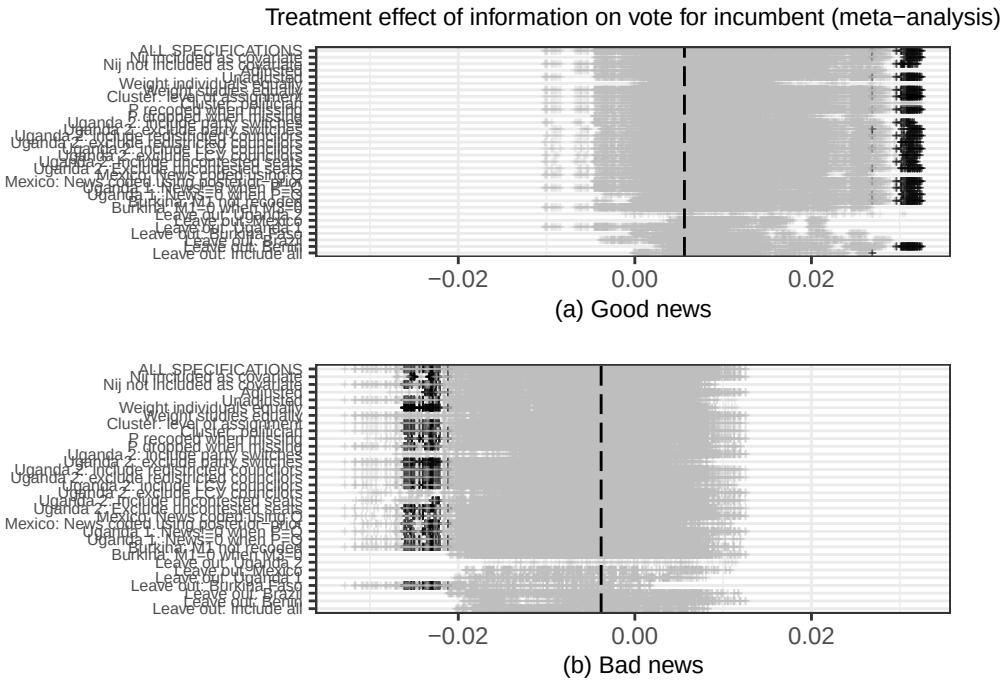

Figure 4: Specification Distribution: Vote for Incumbent. Distribution of average treatment effects on vote for incumbent for a given specification choice, varying all other choices. Darkened vertical lines show estimates for which $p < 0.05$. The dashed vertical line indicates average treatment effect reported in Table 11.1.

5 Figures 4 and 5: Robustness of Findings Across Specifications

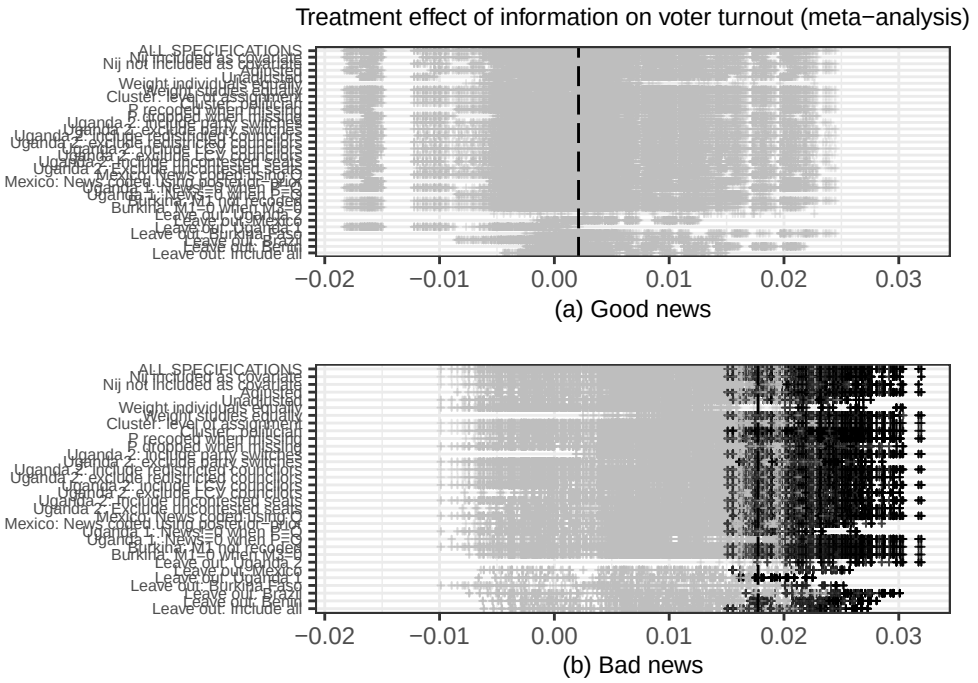

Figure 5: Specification Distribution: Turnout. Distribution of average treatment effects on voter turnout for a given specification choice, varying all other choices. Darkened vertical lines show estimates for which $p < 0.05$. The dashed vertical line indicates average treatment effect reported in Table 11.1.