Making a Difference: The Consequences of Electoral Experiments

Replication Package

Tara Slough

December 10, 2023

README

Overview

The code in this replication package constructs the analysis file from the multiple data sources using R. Five main files run all of the code to generate the data and outputs for the 13 figures and 2 tables in the paper and supplementary materials that rely on data. The replicator should expect the code to run for less than 10 minutes.

This replication package *does not* contain code for the following tables and figures because they do not rely on original analysis of data: Tables 1, 2, 3, A1, A3, A4, A5, A7, A8, A9, and Figure A1. Note further that the output for Table A2 replicates only the right columns of the table that rely on data.

Data Availability and Provenance Statements

Code for data cleaning and analysis is provided as part of the replication package. It is available at https://doi.org/10.24433/CO.7729631.v1.

Statement about Rights

I certify that the author of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

- All data **are** publicly available.

Details on each Data Source

The data come from multiple sources, as enumerated below.

1. **Colorado electoral data**
    - Electoral data is downloaded from the Colorado Secretary of State's databases of election results and voter registration.
        - Precinct-level electoral results are downloaded from https://www.coloradosos.gov/pubs/elections/Results/archive2000.html. The associated raw data files are:
            * `data/raw/2010GeneralPrecinctResults.xlsx`
            * `data/raw/2012GeneralPrecinctLevel.xlsx`
            * `data/raw/2014GeneralPrecinctResults.xlsx`
            * `data/raw/2016_General_Election_Precinct-Level_Results.csv`

    * `data/raw/2018GEPrecinctLevelResults.xlsx`
  - Precinct-level partisan voter registration data are downloaded from https://www.sos.state.co.us/pubs/elections/VoterRegNumbers/2018VoterRegNumbers.html. The raw data is:
    * `data/raw/co2018_reg.csv` (aggregated at precinct level, November 2018)
    * `data/raw/partian_reg_sh.xlsx` (aggregated at level of state house district, 2010-2014)
    * `data/raw/partisan_reg_sh_2016_on.xlsx` (aggregated at level of state house district, 2016-2018)

  The election predictions for 2018 US House races come from the Morris (2018) mid-term election forecasts. They were downloaded in November 2019 from the now-defunct "The Crosstab" website. The raw data associated with these predictions is :
  - `data/raw/indiv_seats.csv`

2. **Pre-analysis plan data**

- Pre-analysis plan and or public pre-registrations were collected in February 2021. I did not access gated/private pre-analysis plans from EGAP or AEA.
  - EGAP pre-analysis plans were provided for all registrations that included as "field experiments" and "elections" through a request to EGAP staff. These were hand-coded coded for mentions of IRB and ethical considerations. The resultant data is included in the following file:
    * `data/raw/egap_merged.csv`
  - The AEA registry contents/metadata was downloaded in February 2021. The current data is subset to experiments that use the keyword "electoral" and adheres to criteria stated in Supplementary Materials section A1.2. These were hand-coded coded for mentions of IRB and ethical considerations. The resultant data is included in the following files:
    * `data/raw/aea_trials_electoral.csv`

3. **Gerber and Green (2000) application**

- This application relies on replication data as well as administrative electoral data from the Connecticut Secretary of State.
  - Replication data from Gerber and Green (2000) was downloaded from https://isps.yale.edu/research/data/d046. The associated raw data file is:
    * `data/raw/NHrep_household.dta`
  - The electoral data from 1998 local and state elections in Connecticut was downloaded from the Connecticut Secretary of State at https://portal.ct.gov/SOTS/Election-Services/Statistics-and-Data/Statistics-and-Data. The raw data is:
    * `data/raw/ct9698.csv`

4. **Boas, Hidalgo, and Melo (2019) application**

- This application relies on replication data as well as administrative electoral data from the Brazilian Tribunal Superior Eleitoral (TSE).
  - Replication data from Boas, Hidalgo, and Melo (2000) was downloaded from https://doi.org/10.7910/DVN/WPVSMH. The associated raw data file is:
    * `data/raw/bhm_2019.csv`
  - The electoral data from the 2016 municipal elections is downloaded from the Brazilian TSE website using the `electionsBR` package in the replication code.

  - In order to merge the replication and TSE data, the code uses a municipal identifier dataset from Peter Johannessen, downloaded from https://www.peterjohannessen.com/data. The data are available in:
    * `data/raw/Brazil+Municipal+Identifiers.csv`

5. **Back-of-the-envelope MAEIs from information experiments**
   - To construct Figure 1, the replication code reads in data from the original back-of-the-envelope calculations described in Table A8.
     - The data used to make this figure are available in:
       * `data/raw/studies_maeis.csv`

Computational requirements

Software Requirements

- R 4.2.1
  - `dplyr_1.1.3`
  - `electionsBR_0.3.2`
  - `estimatr_1.0.0`
  - `foreign_0.8-82`
  - `ggplot2_3.4.2`
  - `gridExtra_2.3`
  - `lubridate_1.8.0`
  - `magrittr_2.0.3`
  - `mvtnorm_1.1-3`
  - `patchwork_1.1.2`
  - `stringr_1.5.0`
  - `openxlsx_4.2.5.2`
  - `tidyr_1.2.0`

Controlled Randomness

- A random seed is set at line 8 of program `03_colorado_sim.R` and at line 7 of `05_existing_applications.R`

Memory, Runtime, Storage Requirements

Summary Approximate time needed to reproduce the analyses on a standard 2022 desktop machine is <10 minutes.

Approximate storage space needed is 25 MB - 250 MB.

Details The code was last run on a **10-core Apple M1 Pro Laptop with MacOS version 13.2.1 with 340GB of free space**.

Description of programs/code

- Programs in `code/01_data_processing.R` generate the analysis datasets `bhm.Rdata`, `colorado.Rdata`, `gg.Rdata`, and `pap_data.Rdata`. The analysis datasets can be found in the `results` folder of the capsule.
- Programs in `code/02_info_exps.R` generate Figure 1 in the main article. The figure can be found in the `results` folder of the capsule.
- Programs in `code/03_colorado_sim.R` generate Figures 2-3 in the main article in Figure A11 in the supplementary materials. The figures can be found in the `results` folder of the capsule.

- Programs in `code/04_pap_analysis` generate Figures A2-A3 and Table A2 in the supplementary materials. The figures and table can be found in the `results` folder of the capsule.
- Programs in `code/05_existing_applications` generate Figures A4-A10 and Table A6 in the supplementary materials. The figures and table can be found in the `results` folder of the capsule.

Instructions to Replicators

- In CodeOcean Capsule, the `Reproducible Run` button will run all programs to generate all results. To run a subset of the programs, remove relevant `Rscript` commands from the `code/.Rproj.user/run` file.
- If you have downloaded the code and data to run in R (outside CodeOcean), adjust file path to the location of the replication files. For a complete replication (including construction of the analysis data), run `01_data_processing.R` prior to any of the other programs.

List of figures, tables and programs

The provided code reproduces selected tables and figures in the paper, as explained.

Figure/Table #	Program	Line Number	Output file
Table 1	n.a. (no data)		
Table 2	n.a. (no data)		
Table 3	n.a. (no data)		
Figure 1	`02_info_exps.R`	8	`results/Figure_1.pdf`
Figure 2	`03_colorado_sim.R`	71	`results/Figure_2.pdf`
Figure 3	`03_colorado_sim.R`	187	`results/Figure_3.pdf`
Table A1	n.a. (no data)		
Table A2	`04_pap_analysis.R`	73	`results/Table_A2.csv`
Table A3	n.a. (no data)		
Table A4	n.a. (no data)		
Table A5	n.a. (no data)		
Table A6	`05_existing_applications.R`	166	`results/Table_A6.tex`
Table A7	n.a. (no data)		
Table A8	n.a. (no data)		
Table A9	n.a. (no data)		
Figure 1	n.a. (no data)		
Figure A2	`04_pap_analysis.R`	38	`results/Figure_A2.pdf`
Figure A3	`04_pap_analysis.R`	59	`results/Figure_A3.pdf`
Figure A4	`05_existing_applications.R`	25	`results/Figure_A4.pdf`
Figure A5	`05_existing_applications.R`	34	`results/Figure_A5.pdf`
Figure A6	`05_existing_applications.R`	50	`results/Figure_A6.pdf`
Figure A7	`05_existing_applications.R`	65	`results/Figure_A7.pdf`
Figure A8	`05_existing_applications.R`	82	`results/Figure_A8.pdf`
Figure A9	`05_existing_applications.R`	136	`results/Figure_A9.pdf`
Figure A10	`05_existing_applications.R`	150	`results/Figure_A10.pdf`
Figure A11	`03_colorado_sim.R`	261	`results/Figure_A11.pdf`

References

Boas, Taylor C.; Hidalgo, F. Daniel; Melo, Marcus A., 2018, "Replication Data for: Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil", https://doi.org/10.7910/DVN/WPVSMH, Harvard Dataverse, V1, UNF:6:6UQpZBKDEfmX7kLg2wU5Zg== [fileUNF]

Gerber, Alan S. and Donald P. Green (2005), "Replication Materials for: Correction to Gerber and Green (2000), Replication of Disputed Findings, and Reply to Imai (2005)," http://hdl.handle.net/10079/1c5b085.

ISPS Data Archive.

Johannessen, Peter (n.d.), "Brazilian Municipal Identifiers." https://www.peterjohannessen.com/s/Brazil-Municipal-Identifiers.csv

---

Acknowledgements

Some content on this page was copied from Hindawi. Other content was adapted from Fort (2016), Supplementary data, with the author's permission.