Reproduction Materials for: Assessing the Validity of Prevalence Estimates in Double List Experiments

Gustavo Diaz
diazg2@mcmaster.ca

February 18, 2023

Reproduction Materials for "Assessing the Validity of Prevalence Estimates in Double List Experiments"

This document contains instructions to reproduce analyses in the main paper and appendix.

Code files are ordered numerically based on the order in which they should run. Data files are not numbered and are required in the same directory for code to work.

Simulation files take a long time to run but they are not necessary to reproduce the results of the paper since the simulation results are included as data files. A minimalist reproduction requires `02_analysis_attn.R` and `07_visualize_sims.R` only.

Manifest

Code files

- `00_README.pdf`: This set of instructions
- `01_preprocess_attn.R`: Recodes `RecodedData.RData` in preparation for analysis. Produces `attn_rep.RData`. This step is optional
- `02_analysis_attn.R`: Reproduces analyses of sections 2 and 3 in the main paper, plus figures D1-2 and tables D1-D3 in the appendix
- `03_dle_sims_deflation.R`: Simulations for the top panels of Figure 2 in the main paper and figures D3-4 in the appendix. Produces `dle_sims_deflation.rda`. This step is optional
- `04_dle_sims_inflation.R`: Simulations for the bottom panels of Figure 2 and figures D3-4. Produces `dle_sims_inflation.rda`. This step is optional
- `05_dle_sims_large.R`: Simulations for Figure D5. Produces `dle_sims_large.rda`. This step is optional
- `06_cost_sim.R`: Simulations for Figure E1. Produces `cost_sim.rda`. This step is optional
- `07_visualize_sims.R`: Reproduces figures 2, D3-5, and E1 as they appear on paper and appendix

Data files

- `RecodedData.RData`: Data from Alvarez et al (2019). "Paying Attention to Inattentive Survey Respondents." *Political Analysis* 27 (2): 145 - 162]. As downloaded from https://doi.org/10.7910/DVN/

TUUYLQ on February 18, 2023

- `attn_rep.RData`: Data for sections 2 and 3 in the main paper, plus figures D1-2 and tables D1-D3 in the appendix
- `dle_sims_deflation.rda`: Simulation output for the top panels of figures 2 and D3-4
- `dle_sims_inflation.rda`: Simulation output for the bottom panels of figures 2 and D3-4
- `dle_sims_large.rda`: Simulation output for Figure D5
- `cost_sim.rda`: Simulation output for Figure E1

Environment

Hardware

OS: Windows 10 Education 64-bit (10.0, Build 19043)
Processor name: AMD Ryzen 3 2200G
Processor speed: 3.5GHz (4 CPUs)
Memory: 16 GB

Software

```
R version 4.2.1 (2022-06-23 ucrt) -- "Funny-Looking Kid"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)
```

Packages

```
future_1.29.0
RItools_0.3-1
RIQITE_1.0 (https://github.com/li-xinran/RIQITE)
coin_1.4-2
survival_3.3-1
ggh4x_0.2.3
DeclareDesign_1.0.0
estimatr_1.0.0
fabricatr_1.0.0
randomizr_0.22.0
data.table_1.14.4
here_1.0.1
forcats_0.5.2
stringr_1.4.1
dplyr_1.0.10
purrr_0.3.4
readr_2.1.3
tidyr_1.2.1
tibble_3.1.8
ggplot2_3.4.0
tidyverse_1.3.2
```

Run time (hh:mm:ss)

`01_preprocess_attn.R`: 00:00:03

`02_analysis_attn.R`: 00:00:09

`03_dle_sims_deflation.R`: 03:28:33

`04_dle_sims_inflation.R`: 03:26:13

`05_dle_sims_large.R`: 11:08:36

`06_cost_sim.R`: 00:55:44

`07_visualize_sims.R`: 00:00:13