Replication Guide
Hierarchically Regularized Entropy Balancing[1]

Yiqing Xu (Stanford) Eddie Yang (UCSD)

Computational platform:

1. Results in the paper are generated from the replication folder using Ubuntu 20.04.2 LTS with x86_64 linux-gnu.
2. The R version is 4.1.2 (2021-11-01)

Installing packages:

1. Results are generated using the following R packages from CRAN: "CBPS" (0.23), "cem" (1.1.29), "devtools" (2.4.3), "dotwhisker" (0.7.4), "estimatr" (0.30.6), "foreign" (0.8-81), "ggplot2" (3.3.5), "glmnet" (4.1-3), "gridExtra" (2.3), "kableExtra" (1.3.4), "MASS" (7.3-54), "tidyverse" (1.3.1), "survey" (4.1-1).
2. Please install the "hbal" package with

   ```
   devtools::install_github("xuyiqing/hbal@f9c21e858df8347bf25c7dc48cab534340b90805")
   ```

   and install "kbal" with

   ```
   devtools::install_github("chadhazlett/KBAL@f9613aa")
   ```

3. For Appendix Table A1 & A2, three additional packages need to be installed. Run the following code:

   ```
   setwd(path_to_replication_folder)
   install.packages('chibal_1.2.1.tar.gz', repos = NULL, type="source")
   install.packages('logbal_1.2.1.tar.gz', repos = NULL, type="source")
   install.packages('l1hbal_1.2.1.tar.gz', repos = NULL, type="source")
   ```

Folders:

Main	code to replicate Fig. 1 & 2 and Appendix Fig. A6-A9
Data	data for the empirical example from Black & Owens (2016)
Appendix	code to replicate tables and all other figures in the Appendix

Replication sequence:

1. Install the necessary packages from CRAN and GitHub. Set the working directory to the replication folder.
2. Run the master script Run.R to replicate all results in the paper.

[1]For questions related to this replication, please contact Eddie Yang (email: z5yang@ucsd.edu).

3. "Main" folder contains files to replicate results in the main text of the paper.
4. "Appendix" folder contains files to replicate results in the appendix of the paper.

Notes:

1. Please make sure to set the working directory to the root replication folder in *R*.
2. The runtime results in Figure 1, Table A3-A5 and Figure A4 may not replicate exactly as they can be subject to the computational power of the machine and the randomness of the simulation process. Replication results may also vary slightly if the operating system and version are different from those used by the authors.
3. All intermediate and output files will be saved to the "Output" subfolders in "Main" and "Appendix". When you run the code, existing files in the these folders will be overwritten, which is normal.
4. For Figure A3 in "Appendix", you can generate the figure directly by running FigureA3.R. To simulate the result from scratch, run the bootstrap and jackknife files first before running FigureA3.R.
5. Figure and Table A3-A5 take a long time to generate. We provide intermediate files and you can directly replicate the tables without running the simulations.

Mapping file names to figures in the paper:

In Paper	File Name
Figure 1	sim_main.png
Figure 2	judge_main.png
Figure A2	penalty.png
Figure A3	qq.png
Figure A4	additional_runtime.png
Figure A5	corr_ipw.png
Figure A6	judge_balance_contending.png
Figure A7	judge_balance_noncontending.png
Figure A8	judge_no_outcome_mod.png
Figure A9	judge_appendix_ori_result.png
Figure A10	lalonde_rep.png

Runtime:

File	Estimated Runtime
Figure1.R	2 hrs
empirical_example.R	30 mins
TabelA1.R	70 mins
TableA2.R	5 hrs
TableA3_5.R	30 hrs
FigureA3.R	1 hr
FigureA4.R	1 min
FigureA5.R	20 mins
TableA6_FigureA2.R	20 mins
TableA7.R	20 mins
TableA8_9.R	1 min