Readme: Visual Inference and Graphical Representation in Regression Discontinuity Designs

January 31, 2023

1 Overview

This replication package requires several different packages for Stata and R, and also uses one MATLAB file. Software versions used are noted below. We are aware of two software packages where using later versions may change output. For `rdrobust`, different versions from what we use may produce slightly different point estimates (in our testing, this did not impact the subsequent analyses, figures, or tables at all), and we include a copy of the version of rdrobust used (7.3.0), as this version predates the package's releases with version control on GitHub. For `RDHonest`, later versions of the package expect data in a different format, and the relevant R program provides the appropriate command to install the version we used from GitHub through the `devtools` package. We provide all data files produced by the build portion of the script, should users wish to focus on the analysis portion.

To run the replication package, see the rd_graph_main.do file in the root directory. Update the specified directories in the header to match your computing environment. Install any required software. Note that if running the build portion of the package, there are one Mathematica and two R scripts referenced in rd_graph_main.do, which must be run manually at the appropriate times.

The notes below document each figure and table in the paper and how it is produced, except for non-statistical output, such as example graphs and screenshots of our experiment.

Data files (found in the data folder) are self-documented with Stata's variable labels. The Pilot and Phase folders contain results from non-expert

trials, and the Expert folders contain results from the corresponding expert trials (see the Experiment Notes section of the readme and the documentation in the full paper and its appendices for additional information). rd_estimates_micro.dta contains all econometric inferences and estimates.

The data/DGP folder contains the 11 DGP specifications and transformed running variables. The simulated_data folder contains the actual simulated datasets from the DGPs used in the graphical trials and on which the different econometric inference procedures are conducted.

On the computing environment used (a shared research computing server with many Intel Xeon E5-4669 CPUs), the build portion of the script runs in approximately three hours. The analysis portion of the script is much faster at under 15 minutes. RAM requirements are minimal.

See the Tables and Figures sections below for information on where and how each table and figure is generated.

2 Software Versions Used

Stata (17.0 2022-06-01)

- rdrobust 7.3.0 (included in rdrobust folder)
- grc1leg 1.0.5
- estout 3.30
- sepscatter 1.1.0
- onewayplot 2.1.3
- moremata 1.0.1/9.2
- car-stata 1.0.0
- cgmreg 3.0.0
- rdmse 2.0

R (4.2.0)

- RDHonest 0.3.2 (we provide the appropriate command to install this version from GitHub with the devtools package)
- zipfR 0.6-70

- Hmisc 4.5-0
- haven 2.4.1
- foreign 0.8-82
- dplyr 1.0.6
- tidyverse 1.3.1
- devtools 2.4.1

Mathematica (12.3.1.0)

3 Experiment Notes

Player treatment group labels (see Table 1 in paper for additional notes)

Phase I (Pilot 1.0) (Phase 1)

Binwidth x Scaling

- A Small bins, normal scale
- B Big bins, normal scale
- C Small bins, big scale
- D Big bins, big scale

Phase II (Phase 2)

Binwidth x spacing

- A Small bins, ES
- B Big bins, ES
- C Small bins, QS
- D Big bins, QS

Phase III (Phase 3)

Smoothed fit lines (nofit/yesfit for included/excluded) x Vertical line at the policy threshold (yesline/noline for included/excluded)

- A nofit, yesline
- B yesfit, yesline
- C nofit, noline

Phase VI (Phase 4)

Binwidth x Fit

- A Small bins, no fit
- B Big bins, no fit
- C Small bins, yes fit
- D Big bins, yes fit

Phase VII (Phase 5)

- A DGP fit by global polynomial, homoskedastic noise term
- B DGP fit by local polynomial, homoskedastic noise term
- C DGP fit by global polynomial, heteroskedastic noise term
- D DGP fit by local polynomial, heteroskedastic noise term

Experts (Except QSS)

- A Vertical line at treatment threshold, no fit lines, small bins, default axis scaling, even spacing

Experts QSS

- B: Vertical line at treatment threshold, no fit lines, big bins, default axis scaling, even spacing

4 Tables

Table 1 (non-statistical table)

Table 2

Filename

- N/A (table created manually from Stata output)

Generated by

- average_error_rates.do

Table 3

Filename

- N/A (table created manually from Stata output)

Generated by

- estimators_v_human_rd_experts_twoway_clustering.do

5 Figures

Figure 1 (non-statistical figure)

Figure 2

Filename

- cef-rdd-combined

Generated by

- dgp_cefs.do

Figure 3 (non-statistical figure)

Figure 4 (non-statistical figure)

Figure 5

Filename

- ROC_phase1
- ROC_phase2
- ROC_phase3
- ROC_phase6

Generated by

- power_functions.do

Figure 6

Filename

- experts_nonexperts_allbins_power_rd

Generated by

- power_analysis_experts_nonexperts.do

Figure 7

Filename

- expert_sample_rankings_rd_overview

Generated by

- expert_pred_pref.do (+/- symbols added manually)

Figure 8

Filename

- pf_comp_2w_all_estimators_expert
- pf_comp_2w_all_sa_estimators_expert

Generated by

- experts_v_human_rd_experts_twoway_clustering.do

Figure 9

Filename

- pf_comp_2w_experts_estimators_prop
- pf_diff_2w_ak_v_ik_expert_prop

Generated by

- estimators_v_human_rd_experts_twoway_clustering

6 Appendix Tables

Table A.1

Filename

- phase1TE_car

Generated by

- table_teffects_car.do

Table A.2

Filename

- phaseIITE_car

Generated by

- table_teffects_car.do

Table A.3

Filename

- phaseIIITE_car

Generated by

- table_teffects_car.do

Table A.4

Filename

- phaseVITE_car

Generated by

- table_teffects_car.do

Table A.5

Filename

- phaseVIITE_car

Generated by

- table_teffects_car.do

Table A.6

Filename

- N/A (table created manually from Stata output)

Generated by

- decision_theory_risks.do

Table A.7

Filename

- N/A (table created manually from Stata output)

Generated by

- decision_theory_risks.do

Table A.8

Filename

- phase1TE_conf_car

Generated by

- tables_confidence_effects_car.do

Table A.9

Filename

- phase2TE_conf_car

Generated by

- tables_confidence_effects_car.do

Table A.10

Filename

- phase3TE_conf_car

Generated by

- tables_confidence_effects_car.do

Table A.11

Filename

- phase6TE_conf_car

Generated by

- tables_confidence_effects_car.do

Table A.12 (non-statistical table)

Table A.13

Filename

- covariate_balI

Generated by

- tables_covariate_balance.do

Table A.14

Filename

- covariate_balII

Generated by

- tables_covariate_balance.do

Table A.15

Filename

- covariate_balIII

Generated by

- tables_covariate_balance.do

Table A.16

Filename

- covariate_balVI

Generated by

- tables_covariate_balance.do

Table A.17

Filename

- covariate_balVII

Generated by

- tables_covariate_balance.do

Table A.18

Filename

- covariate_prediction

Generated by

- tables_covariate_prediction.do

Table A.19

Filename

- table_dgp_prediction_power

Generated by

- tables_dgp_prediction.do

Table A.20

Filename

- table_dgp_prediction_size

Generated by

- tables_dgp_prediction.do

Table A.21

Filename

- N/A (table created manually from Stata output)

Generated by

- prior_analysis_all_phases_nickell_inferences.do

7 Appendix Figures

Figure A.1

Filename

- ROC_phase1_byDGP
- ROC_phase2_byDGP
- ROC_phase3_byDGP
- ROC_phase6_byDGP

Generated by

- power_functions.do

Figure A.2

Filename

- subjective_prob_correct_phase1
- subjective_prob_correct_phase2
- subjective_prob_correct_phase3
- subjective_prob_correct_phase6

Generated by

- confidence_analysis_figs.do

Figure A.3

Filename

- ROC_phase6_DGP9

Generated by

- power_functions.do

Figure A.4

Filename

- pf_comp_2w_CCT_c_cer_expert
- pf_comp_2w_CCT_bc_cer_expert

Generated by

- estimators_v_human_rd_experts_twoway_clustering

Figure A.5

Filename

- pf_comp_2w_theoretical_estimators_expert_cct_ik_ak

Generated by

- estimators_v_human_rd_experts_twoway_clustering

Figure A.6

Filename

- experts_vs_estimators_rmse2
- experts_vs_estimators_mse_decomposition_alt2

Generated by

- experts_vs_estimators_mse_decomposition_alt2
- experts_vs_estimators_rmse2

Figure A.7 (non-statistical figure)

Figure A.8 (non-statistical figure)

Figure A.9 (non-statistical figure)

Figure A.10 (non-statistical figure)

Figure A.11

Filename

- ROC_phase7

Generated by

- power_functions.do

Figure A.12 (non-statistical figure)

Figure A.13 (non-statistical figure)

Figure A.14 (non-statistical figure)

Figure A.15 (non-statistical figure)

Figure A.16 (non-statistical figure)

Figure A.17 (non-statistical figure)

Figure A.18 (non-statistical figure)

Figure A.19 (non-statistical figure)

Figure A.20 (non-statistical figure)

Figure A.21 (non-statistical figure)

Figure A.22 (non-statistical figure)

Figure A.23

Filename

- rdd_dgp_histograms

Generated by

- dgp_histograms.do

Figure A.24

Filename

- ROC_phase6_symmetric_supports
- ROC_phase6_asymmetric_supports

Generated by

- power_functions.do

Figure A.25

Filename

- es_vs_qs_small_bins_gini
- es_vs_qs_big_bins_gini

Generated by

- es_vs_qs_gini.do

Figure A.26 (non-statistical figure)

Figure A.27

Filename

- response_by_gorder_PhaseI
- response_by_gorder_PhaseII
- response_by_gorder_PhaseIII
- response_by_gorder_PhaseVI

Generated by

- prior_analysis_all_phases.do

Figure A.28

Filename

- ROC_rescaled_phase1
- ROC_rescaled_phase2
- ROC_rescaled_phase3
- ROC_rescaled_phase6

Generated by

- power_functions.do

Figure A.29

Filename

- ROC_phase1_rescaled_byDGP
- ROC_phase2_rescaled_byDGP
- ROC_phase3_rescaled_byDGP
- ROC_phase6_rescaled_byDGP

Generated by

- power_functions.do

Figure A.30

Filename

- pf_diff_expert_nonexpert_rd
- pf_diff_expert_qss_nonexpert_bigbins_rd

Generated by

- estimators_v_human_rd_experts_twoway_clustering

Figure A.31

Filename

- pf_diff_2w_pq_expert
- pf_diff_2w_ik_expert
- pf_diff_2w_bc_expert
- pf_diff_2w_rdhon_t_rot_expert

Generated by

- estimators_v_human_rd_experts_twoway_clustering.do

Figure A.32

Filename

- pf_diff_2w_ik_sa_expert
- pf_diff_2w_bc_sa_expert

Generated by

- estimators_v_human_rd_experts_twoway_clustering.do