The Power of Tests for Detecting p-Hacking

Replication package

Graham Elliott
grelliott@ucsd.edu

Nikolay Kudrin
n.kudrin@queensu.ca

Kaspar Wüthrich
kasparwu@umich.edu

This replication package contains the code for replicating all the results in the paper.

Disclaimer

This software is provided "as is" without warranty of any kind, expressed or implied. For questions and error reports, please contact the authors.

Used Software

- R version 4.3.2
- Matlab R2023b

Content

The full replication package, including all intermediate results and output, can be downloaded from Replication Package on Google Drive. The core of this replication package, which is sufficient to generate all other folders and output files and available in the GitHub repository, consists of:

- 6 main files (3 Matlab files and 3 R files) that should be executed in the following order:
  1. `Step1_Methods_Matter_Data_Work.R`
     - Uses data collected by (Brodeur, Cook, and Heyes, 2022)[1] (the relevant files are in `MethodsMatterData`; see the original data for details) to calibrate the distribution of true effects to the RCT subsample of the data, assuming h follows a $\Gamma(\alpha, \beta)$ distribution with shape parameter α and scale parameter β (Results: $\hat{\alpha} = 0.8547432, \hat{\beta} = 1.8690772$).
     - Replicates the results in Table 3 (empirical application). These are saved in `csvFiles/Empirical_Application_Table3a.csv` and `csvFiles/Empirical_Application_Table3b.csv`.
  2. `Step2_Pseudo_Data_Generation.m`
     - Generates pseudo-random data and statistics (in-sample bias, standard error estimates, first-stage F-statistics, BIC-selected lag) for 1 million studies according to the data-generating processes described in Section 5.1. These statistics allow for quick generation of p-curves for any chosen distribution of effects.
     - All statistics are saved in `Raw_Pseudo_Data/MC_raw_pseudo_data.mat`.
  3. `Step3_Generate_Distributions.m`
     - Generates numerical p-curves based on the statistics from the previous step and according to the data-generating processes described in Section 5.1.

[1] Brodeur, Abel, Cook, Nikolai, and Heyes, Anthony. Data and Code for: Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics. Nashville, TN: American Economic Association [publisher], 2022. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2022-02-16. https://doi.org/10.3886/E120246V2

- The p-curves are saved as `csvFiles/Distributions/A_B_CD_Esided_F.csv`, where:
  - `A` $\in$ {`P0` (null), `P1` (p-hacked thresholding), `P1min` (p-hacked minimum)}
  - `B` $\in$ {`Covariate`, `IV`, `IVF`, `LagLength`, `Cluster`} corresponds to covariate selection, IV selection, IV selection with first-stage F-statistic screening, lag-length selection, and cluster level selection, respectively.
  - `C` $\in \{0, 1, 2, 3\}$, where `0`, `1`, and `2` correspond to $h = 0, 1, 2$, respectively, and 3 refers to the case when h is distributed according to $\widehat{\Pi} = \Gamma(\hat{\alpha}, \hat{\beta})$.
  - `D` $\in \{3, 5, 7\}$ corresponds to the number of controls or instruments in covariate or IV/IVF selection examples, respectively. This field is left empty for lag-length and cluster selection examples.
  - `E` $\in \{1, 2\}$ indicates whether researchers use 1- or 2-sided tests.
  - `F` $\in \{0, 1\}$, where `0` denotes a specific-to-general specification search and `1` denotes a general-to-specific search. This field is set to `1` for lag-length and cluster selection examples.
- Example: `P1_Covariate_15_2sided_1` is a p-curve under the thresholding approach to p-hacking in a covariate selection example with $h = 1$, $K = 5$, 2-sided tests, and a general-to-specific specification search.
- Generates components of Figures 18–22. Each subfigure is saved as `Figures/Distributions/P_B_CD_Esided_F.pdf`, where B, C, D, E, and F are as described above.
- Additionally, the second part of the code generates and saves p-curves for $h \in \{0.25, 0.50, \ldots, 4.5\}$ and the corresponding values of Relative Bias implied by p-hacking. Results are saved in `csvFiles/forAppendixI/Distributions/` and `csvFiles/forAppendixI/`, respectively.

4. `Step4_MC_power_main.R`
   - Runs Monte Carlo simulations to evaluate the power of tests for p-hacking examples described in Sections 5.1 and 5.2.
   - Saves results in `csvFiles/Power_Calculations/RejectionRates_main.csv`.
5. `Step5_MC_power_extensions.R`
   - Runs additional Monte Carlo simulations for extensions considered in Appendices F–I.
   - Saves results in `csvFiles/Power_Calculations/`:
     - `RejectionRates_different_rho.csv`
     - `RejectionRates_different_pmis.csv`
     - `RejectionRates_different_N.csv`
     - `RejectionRates_many_h.csv`
6. `Step6_Figures_and_Tables.m`
   - Generates and saves all remaining figures of the paper. Figures are saved in the corresponding subfolders of `Figures/`.
   - Generates LaTeX code for Tables 2–4, saved in `LaTeX_Tables/`, and CSV versions of Tables 2 and 4, saved in `csvFiles/Power_Calculations/`.

- Folder `Functions`, which contains:
  - `RegSpecifications.m` — computes in-sample bias and standard errors for all combinations of controls in covariate selection [used in Step 2]
  - `IVSpecifications.m` — computes in-sample bias, standard errors, and first-stage F-statistics for all IV specifications [used in Step 2]
  - `LagsBIC.m` — selects the optimal Newey-West lag length via BIC [used in Step 2]

  - NeweyWest.m — computes Newey-West standard errors [used in Step 2]
  - NullAndAlternativeDistributions.m — generates null and p-hacked distributions and corresponding biases for covariate and IV selection examples [used in Step 3]
  - NullAndAlternativeDistributions_var_laglength.m — generates null and p-hacked distributions and biases for lag-length selection [used in Step 3]
  - NullAndAlternativeDistributions_var_clust.m — generates null and p-hacked distributions and biases for cluster-level selection [used in Step 3]
  - save_and_plot_p_distributions.m — saves CSV files with numerical p-curves and creates corresponding figures (Figures 18–22) [used in Step 3]
  - MC_Tests.R — contains tests for p-hacking as functions [used in Steps 1, 4, and 5]
  - MC_power.R — computes power of p-hacking tests given null and p-hacked distributions [used in Step 4]
  - MC_power_extensions.R — computes power of p-hacking tests for the extensions in Appendices F–I [used in Step 5]

- Folder Scripts, which contains:
  - Figure_j.m — generates and saves Figure j, where j = 1,2,3,7,...,17 [used in Step 6]
  - Figures_Power_main.m — generates and saves Figures 4–6 and 27–31 [used in Step 6]
  - Figures_Appendix_FGH.m — generates and saves Figures 23–25 [used in Step 6]
  - Figures_Appendix_I.m — generates and saves Figure 26 [used in Step 6]
  - Table_Empirical_Application.m — generates LaTeX code for Table 3 [used in Step 6]
  - Tables_Publication_Bias.m — generates LaTeX code for Tables 2 and 4 [used in Step 6]

Each file includes additional details and descriptions regarding its content. Additionally, we include the folder MethodsMatterData, which contains the data collected by Brodeur, Cook, and Heyes (2022)[1], which we reanalyze in our empirical application.

All other files and folders in the package can be created by running the above code. The folder structure is shown below. In addition, Table 1 below provides an overview of replication files and outputs.

- csvFiles/
  - Distributions/
    - ClusterSelection/ [e.g., P0_Cluster_1_2sided_1.csv]
    - CovariateSelection/ [e.g., P0_Covariate_13_2sided_1.csv]
    - IVSelection/ [e.g., P0_IV_13_2sided_1.csv]
    - LagLengthSelection/ [e.g., P0_LagLength_1_2sided_1.csv]
  - Empirical_Application/
    - Table3a.csv
    - Table3b.csv
  - ForAppendixI/
    - Distributions/
      - CovariateSelection/ [e.g., P0_Covariate_1_2sided_1.csv]
    - Relative_Bias.csv
  - Power_Calculations/

    - Table2.csv, Table4.csv
    - RejectionRates_different_N.csv
    - RejectionRates_different_pmis.csv
    - RejectionRates_different_rho.csv
    - RejectionRates_main.csv
    - RejectionRates_many_h.csv
- Figures/
  - Distributions/
    - ClusterSelection/ [e.g., P_Cluster_1_2sided_1.pdf]
    - CovariateSelection/ [e.g., P_Covariate_13_2sided_1.pdf]
    - IVSelection/ [e.g., P_IV_13_2sided_1.pdf]
    - LagLengthSelection/ [e.g., P_LagLength_1_2sided_1.pdf]
  - Extensions/
    - Appendix_F/ [e.g., Appendix_F_0_Minimum.pdf]
    - Appendix_G/ [e.g., Appendix_G_0_Minimum.pdf]
    - Appendix_H/ [e.g., Appendix_H_0_Minimum.pdf]
    - Appendix_I/
      - PowerBias_Minimum.pdf
      - PowerBias_Thresholding.pdf
  - Figures_analytical_examples/
    - one_sided/
      - Fig1a.pdf, Fig1b.pdf, Fig1c.pdf
      - Fig2a.pdf, Fig2b.pdf, Fig2c.pdf
      - Fig3a.pdf, Fig3b.pdf
      - Fig7.pdf
      - Fig8a.pdf, Fig8b.pdf
      - Fig9a.pdf, Fig9b.pdf
      - Fig10a.pdf, Fig10b.pdf
      - Fig11a.pdf, Fig11b.pdf
      - Fig12.pdf
    - two_sided/
      - Fig13a.pdf, Fig13b.pdf, Fig13c.pdf
      - Fig14a.pdf, Fig14b.pdf, Fig14c.pdf
      - Fig15.pdf
      - Fig16a.pdf, Fig16b.pdf
      - Fig17a.pdf, Fig17b.pdf
  - PowerCurves/
    - ClusterSelection/ [e.g., Cluster_1_2sided_1_Minimum.pdf]
    - CovariateSelection/ [e.g., Covariate_13_2sided_1_Minimum.pdf]
    - IVSelection/ [e.g., IV_13_2sided_1_Minimum.pdf]
    - LagLengthSelection/ [e.g., LagLength_1_2sided_1_Minimum.pdf]
- LaTeX_Tables/
  - Table_2.tex, Table_3.tex, Table_4.tex
- Raw_Pseudo_Data/
  - MC_raw_pseudo_data.mat

Table 1: Overview of Replication Files and Outputs

Figures / Tables	**Data Created**	**Data Used**	**Replicating File(s)**
	Table3a.csv; Table3b.csv	MethodsMatterData/MM Data.dta; MethodsMatterData/Star Wars Data.dta	Step1_Methods_Matter_Data_Work.R, MC_Tests.R
	MC_raw_pseudo_data.mat		Step2_Pseudo_Data_Generation.m, RegSpecifications.m, IVSpecifications.m, NeweyWest.m, LagsBIC.m
Figures 18-22	Main numerical p-curves (P0_Covariate_03_2sided_1.csv and alike); Relative_Bias.csv; numerical p-curves for Appendix I (P0_Covariate_j_3_2sided_1.csv with $\mathtt{j} = 1, \ldots, 18$)	MC_raw_pseudo_data.mat	Step3_Generate_Distributions.m, NullAndAlternativeDistributions.m, NullAndAlternativeDistributions_var_laglength.m, NullAndAlternativeDistributions_var_clust.m, save_and_plot_p_distributions.m
	RejectionRates_power_main.csv	P0_Covariate_03_2sided_1.csv and alike	Step4_MC_power_main.R, MC_power.R, MC_Tests.R
	RejectionRates_different_rho.csv RejectionRates_different_pmis.csv RejectionRates_different_N.csv RejectionRates_many_h.csv	P0_Covariate_03_2sided_1.csv and alike; P0_Covariate_j_3_2sided_1.csv with $\mathtt{j} = 1, \ldots, 18$	Step5_MC_power_extensions.R, MC_power_extensions.R, MC_Tests.R
Figures 1-3, 7-17			Step6_Figures_and_Tables.m, Figure_j.m, where $\mathtt{j} \in \{1, 2, 3, 7, \ldots, 17\}$
Figures 4-6, 27-31		RejectionRates_power_main.csv	Step6_Figures_and_Tables.m, Figures_Power_main.m
Tables 2 and 4	Table_2.tex; Table_4.tex; Table2.csv; Table4.csv	RejectionRates_power_main.csv	Step6_Figures_and_Tables.m, Tables_Publication_Bias.m
Table 3	Table_3.tex	Table3a.csv; Table3b.csv	Step6_Figures_and_Tables.m, Table_Empirical_Application.m
Figures 23-25		RejectionRates_different_rho.csv, RejectionRates_different_pmis.csv, RejectionRates_different_N.csv	Step6_Figures_and_Tables.m, Figures_Appendix_FGH.m
Figure 26		RejectionRates_many_h.csv; Relative_Bias.csv	Step6_Figures_and_Tables.m, Figures_Appendix_I.m

Data Acknowledgment

This replication package includes data from: **"Data and Code for: Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics"**, by *Abel Brodeur, Nikolai Cook, and Anthony Heyes.*

Available at https://www.openicpsr.org/openicpsr/project/120246/version/V2/view.

This dataset is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license.