

README for “Experimental evidence on the acceptance of males falling behind.”

Authors: Alexander W. Cappelen, Ranveig Falch and Bertil Tungodden.

Overview

The code in this replication package constructs the analysis file from nine data sets which includes experimental data, using Stata and R. Three main files run all of the code to generate the data for the 11 figures and 27 data-driven tables in the paper and the online appendix. The replicator should expect the code to run for less than 1 hour.

Data Availability and Provenance Statements

The data for the main choice experiment, the follow-up choice experiment, and the survey experiment were collected through three survey providers in a series of data collections. The survey providers, Kantar TNS, Ipsos and NORSTAT, are all international data collection agencies recruiting participants through large, existing panels with the aim of recruiting nationally representative samples on a set of observable characteristics (age, gender and broad geography). The use of different survey providers in the different rounds of data collection was a result of the legal requirement for competitive tenders as a public institution in Norway. The data was collected in seven rounds.

Statement about Rights

- ✓ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ✓ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

License for Data

The data are licensed under CC0.

Summary of Availability

- ✓ All data are publicly available.

Details on each Data Source

The experimental and survey data used to support the findings of this study have been deposited in the Harvard Dataverse [Cappelen, Falch, and Tungodden, 2024]. This includes a description of each data set, study documentation as well as instructions in English. The experimental and survey data were collected by

the authors with support from three data-collection agencies, Kantar TNS, Ipsos and NORSTAT and are available under a Creative Commons Non-commercial license. The nine data sets can be downloaded from the Harvard Dataverse Cappelen et al. [2024]. A detailed list is provided in the table below:

Data files				
Data description	Year	Name	Source	doi
Spectator choices	2015	choiceexp_study1r1_2015RAW.dta choiceexp_study1r1_2015RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Spectator choices	2016	choiceexp_study1r2_2016RAW.dta choiceexp_study1r2_2016RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Spectator choices	2020	choicexp_study2_2020RAW.dta choicexp_study2_2020RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Beliefs, choice and survey experiment	2021	surveyexp_round2_Sept2021RAW.dta surveyexp_round2_Sept2021RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Follow-up, beliefs and spectator choice	2024	followup_2024RAW.dta followup_2024RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Follow-up, hand categorization	2024	handcategorization.dta handcategorization.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Follow-up, API categorization	2024	trainedgptcategorization.dta trainedgptcategorization.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Government support	2022	surveyexp_round3_July2022RAW.dta surveyexp_round3_July2022RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G
Effort beliefs, survey experiment	2019	surveyexp_round1_Dec2019RAW.dta surveyexp_round1_Dec2019RAW.tab	Harvard Dataverse	10.7910/DVN/J0XC2G

In addition, to facilitate seamless comparison between the tables in the paper and the code, we have created a separate data file, `US_stats.dta`, containing a few descriptive statistics for the US population. This data file is in `replication-package\estimations\data\raw`.

Computational requirements

Software Requirements

- Stata (code was last run with version 18.5, but with version specifications in the different do-files).
 - `estout`, version 3.31
 - `coefplot`, version 1.8.6
 - `lincomest`, version 10.0
 - R 4.3.1
 - `haven`, version 2.5.4
 - `magrittr`, version 2.0.3
 - `Adjust.Rdata`
(found in the replication package in `estimations\scripts\Adjust_R_code`).
- Our computation of p-values adjusted for step-down multiple testing follows the algorithm proposed by Romano and Wolf [2016].

`estimations\scripts\Adjust_R_code` contains the code (and documentation) of the Romano-Wolf procedure, which was kindly made available by Michael Wolf (`Adjust.Rdata` and `ReadMe.txt`) in public domain.

Controlled Randomness

- Random seed for the multiple hypotheses testing is set at the following lines in each, respective do-file for multiple hypotheses testing:
 - Line 175, 311 and 447 in `pvalues_adjustment_choiceexp.do`
 - Line 242, 377 and 513 in `pvalues_adjustment_surveyexp2.do`
 - Line 195 and 332 in `pvalues_adjustment_surveyexp3.do`
 - Line 193 and 328 in `pvalues_adjustment_surveyexp_s1s2merged.do`

Memory and Runtime Requirements

The code was last run with Stata and RStudio on a Lenovo ThinkPad T480s 4-core Intel Core i5-based laptop with Windows. Approximate time needed to reproduce the analyses on a standard 2024 desktop machine version 22H2.

✓ 10-60 minutes

Description of Code and Instructions to Replicators

- Download the nine data sets in Cappelen et al. [2024] in .dta format and store them in `replication-package\estimations\data\raw`
- Run `master.do` in `replication-package\estimations\scripts`.
 - This do-file runs seven cleaning do-files;
`choiceexp_study1r2_cleaningformerge.do`,
`choiceexp_study1_cleaning.do`,
`surveyexp_round2_cleaningformerge_wchoice.do`,
`choiceexp_study2_cleaning.do`,
`surveyexp_round2_cleaningformerge_wr1.do`,
`surveyexp_round3_cleaning.do`,
`choiceexp_followup2024_cleaning.do`; which cleans the data sets.
 - It then runs eight do-files which reproduces tables, figures and numbers, referencing the place to find each of them in the article; `choiceexp.do`, `surveyexp_round1.do`, `surveyexp_round2.do`, `surveyexp_round3.do`, `surveyexp_merged_r1r2.do`, `pooling_survey123_fordescriptives.do`, `followup2024.do` and `US_stats.do`. These do-files generate all tables, figures and numbers in the main body of the article and in the online appendix (except for the multiple hypotheses tests and illustrations not based on data). The file `Overview.xlsx` provides an overview of all figures and tables, which do-file generates each of them, and on what line in the do-file they are saved (`replication-package\estimations\output`). Note

- that some esthetic changes has been made to the figures using the graph editor manually, such as changes in label sizes.
- Finally, it runs the four do-files which complete the STATA-part of the multiple hypotheses testing; `pvalues_adjustment_choiceexp.do`, `pvalues_adjustment_surveyexp2.do`, `pvalues_adjustment_surveyexp3.do` and `pvalues_adjustment_surveyexp_s1s2merged.do`, storing twenty .dta-files for the next step of the multiple hypotheses testing in `replication-package\estimations\data\processed`, and excel tables with coefficients from each multiple hypothesis test: `tableA14.xls – tableA23.xls` in `replication-package\estimations\output\tables`.
 - The next step for the multiple hypotheses testing is to run `mht.R` in R. The adjusted p-values are stored in `replication-package\estimations\data\processed`, one file per table: `pvals_adj_tableA14.dta – pvals_adj_tableA23.dta`.
 - Run the Stata do-file `pvals_release.do` in `replication-package\estimations\scripts` to add the raw and adjusted p-values to each multiple hypothesis table `tableA14.xls – tableA23.xls`.
 - Note the following slight discrepancies between the tables generated by the code and those presented in the text, primarily due to manual adaptations:
 - Column names produced by the code closely align with those in the text but are not exact matches.
 - Commas have been applied as thousand separators in the tables, particularly for the number of observations and income values.
 - The “Prob > F” value is included as a note in each relevant table rather than as a separate column.
 - The stars in the last column of Tables A7, A8, and A9 are manually added based on regression results, due to limitations in `esttab` when working with `lincom` in STATA.
 - Figures 3 and 6 appear in multiple colors in the text, but may appear grey in the code outputs.

References

- Alexander W. Cappelen, Ranveig Falch, and Bertil Tungodden. Replication Data for: Experimental Evidence on the Acceptance of Males Falling Behind, 2024. URL <https://doi.org/10.7910/DVN/J0XC2G>.
- Joseph P. Romano and Michael Wolf. Efficient computation of adjusted p -values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, 113(1):38–40, 2016. doi: 10.1016/j.spl.2016.02.012.