Data and Code for Replication of "Checking and Sharing Alt-Facts"

by Emeric Henry, Ekaterina Zhuravskaya, Sergei Guriev (AEJ:Policy)

The data citation and access conditions:

- The paper uses original data created by authors:
  Henry, Emeric, Ekaterina Zhuravskaya, and Sergei Guriev. 2020. "Data and Code for: Checking and Sharing Alt-Facts" *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E140161V1.
- All data are publicly available.

Software requirements:

- Stata 16.1.

- Additional Stata packages: cibar, grc1leg, estout, outreg2, and st0085_2. They are installed automatically (if not already pre-installed) by **Master_replication_HZG.do.**

- The replication was run on Stata/MP 16.1 for Mac (Intel 64-bit).

Computational requirements:

- Minor: The size of the dataset is 3.3MB. It has 246 variables and 5,089 observations.

Time requirements:

- Less than 10 minutes.

Instructions for replication:

To generate all results in the article: Quit and Relaunch Stata 16.1, run the **Master_replication_HZG.do**, it first calls the dofile that combines and cleans the original data and then calls the dofile which generates all the tables and figures in the paper, then calls the dofile that generates additional numbers used in the text. See Tables A and B in this readme file for a guide on which parts of the code replicate each result in the article).

To run this file, define the path to the replication folder by replacing the line 13 in this dofile with the actual path on your computer. (All other paths in the code are relative to $dir).

The structure of the replication folder:

Root directory "replication" should contain the dofile **Master_replication_HZG.do** and the following subfolders:

- **original_data**

- **dofiles**

Master_replication_HZG.do also generates the following subfolders (if they do not already exist) in the root directory "replication":

- **generated_data**

- **results**

The contents of the subfolders:

original_data subfolder contains:

1) File "Description and codebook of original data files"

2) The original data files:
   – Survey+1_May+27%2C+2019_01.02.csv
   – Survey+2_May+27%2C+2019_01.03.csv
   – Survey+3_May+27%2C+2019_01.03.csv
   – Survey+4%2C+5+and+6+-+RANDOMIZED_May+27%2C+2019_01.04.csv
   – share_facebook.dta
   – GA_hours.dta

3) The Online Appendix for the paper, which contains the codebook for the survey data: ONLINE_APPENDIX_HZG_final_AEJ_Pol.pdf (see Online Appendices C and D).

dofiles subfolder contains two dofiles which are called by **Master_replication_HZG.do** in the following order:
- **"1.infile_data.do"** (which infiles and cleans the original data and generates the file **surveys.dta** in the **generated_data** subfolder)
- **"2.analysis_final_paper.do"** (which uses **surveys.dta** to generate the tables and figures, which are placed in the **results** subfolder).
- **"3.numbers_in_text.do"** (it generates additional numbers mentioned in the text and places then in the log file **numbers_in_text.smcl** saved in the **results** subfolder).

generated_data subfolder is generated (or re-written) by **Master_replication_HZG.do**, which calls **1.infile_data.do** to create surveys.dta in this subfolder.

results subfolder is generated (or re-written) by **Master_replication_HZG.do**, which calls **2.analysis_final_paper.do** to generate all the results (tables and figures) in the paper (including online appendix tables and figures) to be placed in this subfolder, and **3.numbers_in_text.do** generates additional numbers mentioned in the text. See also Table B in this readme file for the provenance of the numbers mentioned in the text.

NB! To replicate the results of the paper exactly, it is important to quit and relaunch Stata before executing Master_replication_HZG.do. This is because a subset of the results is generated by LASSO, which requires a random number as an entry. Despite the "set seed", if the **Master_replication_HZG.do** run several times without quitting the Stata, the LASSO results are slightly different. (Different versions of Stata could also generate the same issue.) Importantly, we have run many tries and even though the LASSO results can be slightly different each time (if Stata is not relaunched), the differences are very small (and affect only

the third or the fourth digit). These changes do not in any way affect the conclusions of the paper.

The randomization defined by set seed affects the following tables:

- Panel C of Table 2. (This panel uses the LASSO prediction to define the dependent variable). The exact LASSO prediction depends on the random seed.

- Table A5. This table uses the covariates selected by LASSO and used for LASSO prediction. The exact LASSO prediction depends on the random seed.

- Tables A6. This table is the detailed output of the regressions presented in Panel C of Table 2.

- Table A7. This table uses the covariates selected by LASSO and used for LASSO prediction. The exact LASSO prediction depends on the random seed.

Table A. The lines of the code for replicating tables and figures:

Table or Figure	#	Program	Line
Table	1	2.analysis_final_paper.do	694-721
Table	2	2.analysis_final_paper.do	763-1070
Table	3	2.analysis_final_paper.do	1075-1119
Table	A1	2.analysis_final_paper.do	726-756
Table	A2	2.analysis_final_paper.do	1187-1189
Table	A3	2.analysis_final_paper.do	765-848
Table	A4	2.analysis_final_paper.do	857-924
Table	A5	2.analysis_final_paper.do	168-376
Table	A6	2.analysis_final_paper.do	932-1069
Table	A7	2.analysis_final_paper.do	383-683
Table	A8	2.analysis_final_paper.do	1130-1138
Figure	2	2.analysis_final_paper.do	40-90
Figure	3	2.analysis_final_paper.do	91-130
Figure	4	2.analysis_final_paper.do	135-165
Figure	5	2.analysis_final_paper.do	168-376
Figure	6	2.analysis_final_paper.do	383-683
Figure	A1	2.analysis_final_paper.do	1198-1209
Figure	A2	2.analysis_final_paper.do	1215-1249

Table B. Replicating numbers cited in the text:

Place in the text of the article:	Number to be replicated:	Formulae, Tables, or the relevant lines in the program 3.numbers_in_text.do
1. Introduction, paragraph starting "All 2,537 participants"	2,537 302 82 21	Lines 6-28 in **3.numbers_in_text.do**
	11.9% 3.2%	Based on numbers above: 11.9%=302/2537

		0.8%	3.2%=82/2537 0.8%=21/2537
2.	Introduction, paragraph starting "The central result"	14.7%	Table 2 Panel A Column 1 "Mean DV …"
		10.2%	=14.7%-0.045 (for 0.045, see Table 2 Panel A Column 1, coefficient on "Imposed Fact-Check")
		10.8%	=14.7%-0.038 (for 0.038, see Table 2 Panel A Column 1, coefficient on "Voluntary Fact-Check")
		More than 25%	Based on numbers above: (14.7-10.2)/14.7=27%; (14.7-10.8)/14.7=31%
		4.6%	=0.0462, see Table 2 Panel A Column 5 "Mean DV …"
		2.6%	=0.0462-0.020 (for 0.020, see Table 2 Panel A Column 5, coefficient on "Imposed Fact-Check"
		2.5%	=0.0462-0.021 (for 0.021, see Table 2 Panel A Column 5, coefficient on "Voluntary Fact-Check")
		43 to 45%	Based on numbers above: 0.020/0.0462=43%; 0.021/0.0462=45%
3.	Abstract	About 45%	See the line just above (43 to 45%)
4.	Introduction, paragraph starting "Those participants"	3.15%	Lines 33-34 in **3.numbers_in_text.do**
		3.07%	Lines 35-36 in **3.numbers_in_text.do**
		14.3%	See Table 2 Panel B Column 1 "Mean DV …"
		11.5%	=0.143-0.028 (for 0.028, see Table 2 Panel B Column 1, coefficient on "Voluntary Fact-Check"
5.	Introduction, paragraph starting "At first glance"	39%	39%=330/846 where 330 is the number of Viewers (see Line 42 in **3.numbers_in_text.do**) and 846 is from Table 1 Column 3 "Observations"

6.	Introduction, paragraph starting "The key to the puzzle"	39%	-39% = -0.023/0.0589
		60.2%	-60.2% =-0.026/0.0432
			For 0.023, 0.0589, 0.026, 0.0432, see Point #18 in this Table B (below)(*the replication for II.D Paragraph starting "Panel C of Table 2"*)
7.	Section I.B	2,537	Table 2 Panel A Column 1, "Observations"
8.	Section I.D	Less than 800	See Figure 2
9.	Section I.E	0.74	Table A1, column 1 at the bottom of the table "p-value: joint significance by treatment"
		0.17	Table A1, column 2 at the bottom of the table "p-value: joint significance by treatment"
		0.26	Table A1, last row, "p-value: joint significance both treatments"
		35 million 27 million	See the reference cited in footnote 13 of the article
		70%	Line 48 in **3.numbers_in_text.do** generates 67.84 which we round up to 70%
		50%	Table 1, simple average of columns (1), (2), (3), "Often share on FB" yields 48%, which we round up to 50%
		214	Table A2, first column, "Number of friends"
		338	Table A2, last column, "Number of friends"
10.	Section II.A, first paragraph	749 612 52	Lines 50-96 in **3.numbers_in_text.do**
		846	Table 1 Column 2 "Observations"
		330	The number of Viewers (see Line 42 in **3.numbers_in_text.do**)
11.	Section II.A, second paragraph	2,537 302 82 21	Same as Point #1 of this Table B (above)

	1,176 218 53 7	Lines 98-119 in **3.numbers_in_text.do**
12. Section II.B	All numbers	Same as Point 2 of this Table B (above) *(Introduction, paragraph starting "The central result")*
13. Section II.C, first paragraph	All numbers	Same as Point 4 of this Table B (above) *(Introduction, paragraph starting "Those participants")*
14. Section II.C, second paragraph	39%	39%=330/846 where 846 is from Table 1 Column 3, "Observations" 330: the number of Viewers (see Line 42 in **3.numbers_in_text.do**)
15. Section II.D, first paragraph	846 330 39%	846: Table 1 Column 3, "Observations" 330: the number of Viewers (see Line 42 in **3.numbers_in_text.do**) 39%=330/846
16. Section II.D, paragraph starting "The following factors"	20 percentage points	Table A5 Column 1, coefficient on "Voted Le Pen, 2nd round 2017"
	7.3 percentage points	Table A5 Column 2, coefficient on "Voted Le Pen, 2nd round 2017"
	8 percentage points	Table A5 Column 1, coefficient on "Feel closer to European gov."
	3.9 percentage points	Table A5 Column 2, coefficient on "Feel closer to European gov."
	6.3 percentage points	0.257 see Table A5 Column 1, coefficient on "Reason to share: influence" times Std. Dev. 0.24 (see Table 1 Column 2 "Reason to share: influence"), 0.257*0.24 up to a rounding error
	1.3 percentage points	0.054 see Table A5 Column 2, coefficient on "Reason to share: influence" times Std. Dev. 0.24 (see Table 1 Column 2 "Reason to share: influence"), 0.054*0.24=1.3

	14.5 percentage points	Table A5 Column 1, coefficient on "Religious"
	5.6 percentage points	Table A5 Column 2 , coefficient on "Religious"
	5.3 percentage points	Table A5 Column 1, coefficient on "Male"
	2.7 percentage points	Table A5 Column 2, coefficient on "Male"
	7.6 percentage points	Table A5 Column 1, coefficient on "Donated blood"
	2.6 percentage points	Table A5 Column 2, coefficient on "Donated blood"
17. Section II.D, paragraph starting "In the charts"	18.8%	Table A6 Column 1 "Mean Predict., Voluntary T, Viewer"
	13.6%	Table A6 Column 1 "Mean Predict., Voluntary T, Nonviewer"
	5.9%	Table A6 Column 5 "Mean Predict., Voluntary T, Viewer"
	4.3%	Table A6 Column 5 "Mean Predict., Voluntary T, Nonviewer"
18. Section II.D, paragraph starting "Panel C of Table 2"	0.023	Table A6 Column 5, coefficient on "Voluntary Fact-Check: Viewer"
	0.0589	Table A6 Column 5 "Mean Predict., Voluntary T, Viewer"
	0.026	Table A6 Column 5, coefficient on "Voluntary Fact-Check: Nonviewer"
	0.0432	Table A6 Column 5 "Mean Predict., Voluntary T, Nonviewer"
	0.066	Table A6 Column 1, coefficient on "Voluntary Fact-Check: Viewer"
	0.136	Table A6 Column 1 "Mean Predict., Voluntary T, Viewer"
	0.025	Table A6 Column 1, coefficient on "Voluntary Fact-Check: Nonviewer"
	0.188	Table A6 Column 1 "Mean Predict., Voluntary T, Nonviewer"
19. Section II.D, paragraph starting "Similarly to the analysis"	2.6 percentage points	Table A5 Column 3, coefficient on "Negative image EU"
	One percentage point	Table A5 Column 4, coefficient on "Negative image EU"

	1.6 percentage points	Standard deviation of self-reported reciprocity: 0.18229503 see line 127-128 in **3.numbers_in_text.do**). 1.6=0.086*0.18229503 For 0.086, see Table A5 Column 3, coefficient on "Self-reported reciprocity"
	0.8 percentage points	Standard deviation of self-reported reciprocity: 0.18229503 see Lines 127-128 in **3.numbers_in_text.do**). 0.8=0.046*0.18229503 For 0.046, see Table A5 Column 4, coefficient on "Self-reported reciprocity"
20. Section II.D, paragraph starting "This is why we reproduce"	15.3% 7.0% 6.3% 4.8% 2.5% 2.1%	Lines 136-155 in **3.numbers_in_text.do**
	27% to 31% 43% to 45%	Same as Point 2 of this Table B (above) *(Introduction, paragraph starting "The central result", sub-point* "More than 25%" *and sub-point* "43 to 45%"
	54% to 59%	Lines 156-157 in **3.numbers_in_text.do**
	47% to 56%	Lines 160-161 in **3.numbers_in_text.do**
21. Section III, Footnote 23	135 190 67	Line 169 in **3.numbers_in_text.do**
22. Section V, Conclusions	43% to 45%	Same as Point 2 of this Table B (above) *(Introduction, paragraph starting "The central result", sub-point* "43 to 45%"