Replication report: Checking and Sharing Alt-Facts*

Monica Beeder Erik Ø. Sørensen

April 24, 2023

Abstract

Henry, Zhuravskaya, and Guriev (2022) examine whether people are willing to share "alternative facts" espoused by right-wing populist parties before the 2019 European elections in France and how this interacted with the availability of fact-checking information. They find that both imposed and voluntary fact-checking reduce the likelihood of sharing false statements by approximately 45%, and that imposed and voluntary fact-checking have similar effect sizes. We reproduce these findings and introduce several alternative estimates to assess the robustness of the original results, including resolving an inconsistency in the handling of pre-treatment controls. Overall, our results align with the results of the original paper. The differences we find are small in absolute magnitude but, since many effects were small, not always trivial in terms of relative differences. This replication supports the conclusions of the original paper.

Keywords: *Replication, Fake News, Alternative facts*

1 Introduction

Henry et al. (2022) investigated the willingness of people to share "alternative facts" propagated by a right-wing populist party prior to the 2019 European elec-

*Both authors: FAIR, Department of Economics, NHH Norwegian School of Economics. Emails: monica.beeder@nhh.no and erik.sorensen@nhh.no. The replication was conducted as part of Oslo Replication Games, organized by The Institute for Replication. A detailed guide to the replication (and all the code) can be found at Zenodo, https://doi.org/10.5281/zenodo.7858828. We are grateful to the original authors for making all the data, code, and their pre-analysis plan available. This replication was supported by the Research Council of Norway through its Centres of Excellence Scheme, FAIR Project 262675.

tions in France, and how this behavior was influenced by the availability of fact-checking information. Henry, Zhuravskaya and Guriev (henceforth referred to as "HZG") conducted a randomized online experiment with 2,537 French voting-age Facebook users. All participants were asked if they wanted to share two "alternative" statements by the populist party on Facebook. The treatment groups received either imposed or voluntary fact-checking information alongside the misleading statements.

One-third of the participants received voluntary fact-checking alongside the false statements, another third received imposed fact-checking with the false statements, and the remaining group only received the false statements. Participants were then asked about their willingness to share the statements on their personal Facebook page, this corresponds to the 'Intent to share alt-facts on Facebook' outcome, for which treatment effects are estimated in the left panels of their Table 2. If participants expressed a desire to share the statement, they were directed to a second window with a button resembling a Facebook "share" button, which they could click on to share the statements. This corresponds to the 'Action of sharing alt-facts on Facebook' outcome, for which treatment effects are estimated in the right panels of their Table 2. When participants clicked the button, the statement was not automatically shared on Facebook; instead they were redirected to a different site displaying how the post would appear on their Facebook page. From there, they had the option of sharing the post on Facebook.

The main conclusions of the original paper were that the fact checking (both voluntary and imposed) reduced the sharing of false statements, and that the distinction between fact checking being imposed or voluntary did not matter much. Furthermore, introducing an extra step to share false statements online reduced the actual sharing substantially.

In our replication, we focused on HZG Table 2, *Average Treatment Effects*, since it presents the main results of the original paper. Additionally, we concentrated on columns 4 and 8, which show the complete specification with all control variables included, for the "intent" and the "action" outcome respectively. We did not replicate the figures, as they are visual representations of results already available in HZG Table 2.

2 The original analysis and replication variants

For each of the two outcomes, "intent" and "action," there are three panels of HZG's Table 2, corresponding to different main results. In Panel A, the sharing outcomes are regressed on treatment dummies for the full sample, providing average treatment effects. In Panel B, the sharing of fact-checking information is regressed on a dummy for whether the fact-check was voluntary (as opposed to

imposed) in the sample of the two treatments in which fact checking was available. In Panel C, the analysis attempts to control for pre-treatment propensity to share alt-facts. A lasso regression is run as a first step to predict the sharing of alternative facts in the treatment without access to fact checking. The estimated lasso model is then used to predict propensity to share alternative facts in the other two treatments, and the second step regression outcome is the difference between the actual and predicted sharing. In this second step regression, a distinction is made among the participants in the voluntary fact check treatment between those that viewed and those that did not view the fact check.

For the replication, our initial goal was to precisely reproduce HZGs original results. We refer to this exercise as "**O**" (for "original"), and use it is a baseline. Initially, we faced difficulty reproducing the original results using minimally modified versions of the original code but using Stata SE v17 instead of Stata SE v16.1 as reported used by HZG. However, locating a computer with an older installation of Stata allowed us to implement **O** exactly, prompting us to include a replication variant "**V**" (for "version") using Stata SE v17.

Upon reviewing the code, we identified a sample restriction that could be removed. HZG only included study participants who spent a minimum of 250 seconds on the online survey. Neither the paper's text nor the pre-analysis plan mentions this restriction. In replication variant "**S**" (for "sample"), we remove this restriction.

The full specification in HZG Table 2, columns 4 and 8, relies on a linear formulation with numerous control variables. While the lasso shrinkage and selection procedure (Tibshirani, 1996) is applied in the prediction of who wants to share information (for the outcomes panel C of HZG Table 2), the second stage estimation of treatment effects does not consider variable selection and functional form specifically.[1]. In replication variant "**L**" (for "lasso"), we broaden the set of control variables by generating a complete set of quadratics and interactions. We use this expanded set of control variables both for the initial prediction step and the subsequent estimation of treatment effects using the Chernozhukov et al. (2018) estimator.

While working with other replication variants, we observed an apparent anomaly in the number of observations reported in HZG Table 2. In panel C, Column 4 reports 2,018 observations, while Column 8 reports 2,078. Considering both regressions are meant to use "all pre-treatment characteristics" in the second step it seemed puzzling that the initial lasso variable selection somehow could select for a smaller sample in Column 4 than in Column 8. Further inspection revealed differences in the co-variates included as control variables in the outcome regressions and the lasso regression for prediction, as well as discrepancies in the handling of

[1]The lasso was not mentioned in the pre-analysis plan

missing observations. As an example, for the lasso prediction step all observations are included whether they did or did not report income, with a dummy introduced indicating missing incomes (124 observations). Conversely, participants who did not respond to a question on voting for Macron in the 2017 first round (108 observations) are excluded from the prediction sample. However, in the treatment regressions (columns 4 and 8 of HZG Table 2), the 124 observations with missing incomes are excluded from the estimation sample and the 108 missing information on voting for Macron are coded as if they did not vote for Macron in the 2017 first round. This appears inconsistent. We suspected the major effect of these relatively minor differences to go through sample selection. In replication variant "**C**" (for "controls") we adopt an inclusive approach for these two variables: First, we control for missing incomes using a dummy variable (as HZG did in the prediction step). Second, we control for missing information on 2017 voting by coding all observations that did not explicitly report voting for Macron in the 2017 first round as if they did not (as HZG did in the second step regression). We also make sure that the same set of variables are used as potential controls in both the prediction step and in the treatment effect estimation.

3 Replication results

Our replication results are provided in two tables, the first one (Table 1) estimating treatment effects on the *intent* to share alt-facts, focussing on Column 4 of HZG Table 2 and the second (Table 2) estimating treatment effects on the *action* to share alt-facts, focussing on Column 8 of HZG Table 2. The first column in our tables is the results as presented in the original paper for comparison (labelled "in *AEJ*"), while the remaining columns report on the replication variants as defined in Section 2.

First, we can note that in both Table 1 and Table 2, there are no differences between the "In *AEJ*," and the **O** replication variant; the results HZG reported reproduce exactly.

In Table 1, there are also no differences between the **O** and the **V** variants, the choice of Stata version has no impact on these results. The same holds in Table 2A and 2B, but we notice that in Table 2C, the number of observations drop by 60 (or about 3%),[2] and this is accompanied by a somewhat stronger negative treatment effect of having imposed Fact Check on the action of sharing alternative facts, a change of about -9%.[3] Given the effect on the sample size, this difference must follow from how Stata 16.1 and Stata 17 implement the lasso selection of control variables in the prediction step. The other differences between the **O** and **V**

[2] From 2,078 to 2,018.

[3] From -0.02094 to -0.02282.

specification in Table 2C are trivially small.

Removing the speeders, as implemented in variant **S**, has only minor effects on both the sample size and the estimated treatment effects, only 12 speeders were removed. It is the case that the relative difference in the effect of voluntary fact check on the viewers intent to share change by about 13%,[4] and the effect of voluntary fact check on action increases by about 27%,[5] but both these effects were (and remain) small and non-significant.

The introduction of the lasso also for the estimation of treatment effects (variant **L**) has a large effect on the number of observations included, apparently a large number of observations were excluded in the original specifications without being very important in the estimation of treatment effects, in Table 1A we see the the number of observations increase by 9%.[6] A number of estimated treatment effects get smaller. The magnitude of the two treatment effects in Table 1A drops by about 20% and 17%, while remaining statistically significant.[7] In Table 1B, the magnitude of the estimated treatment effect drops by 36% and is no longer (marginally) significant.[8] The relative differences are also large in Table 2B, but the absolute magnitudes are very small.

Consistent treatment of control variables across the prediction and the treatment estimation step (variant **C**) was constructed to increase the total sample size, and we see that it increases by about 3%.[9] The effects on the estimated treatment effects are mostly modest. The magnitude of the negative effect of imposed fact checks in Table 1A decrease by about 10%,[10] and the magnitude of the negative effect of imposed fact checks in Table 1B increases by about 10%;[11] other differences are smaller in relative terms or too close to zero to have any major impact.

4 Concluding remarks

We confirm that the original analysis reproduces exactly. Having also conducted several robustness checks, we conclude that qualitatively, the results of the original replicate. When our robustness checks result in substantial differences in the relative size of estimated coefficients, this occurs when the absolute magnitude of the effects is small, rendering the relative differences less important. This is the case even for a robustness variant that resolves an apparent inconsistency in the

[4] From −0.02538 to −0.02199

[5] From 0.00368 to 0.00468.

[6] From 2,078 to 2,266.

[7] From −0.04999 to −0.04011 and from −0.05202 to −0.04341.

[8] From −0.03896 to −0.02510.

[9] From 2,078 to 2,146.

[10] From −0.04999 to −0.04499.

[11] From −0.03896 to −0.04284.

treatment of pre-treatment characteristics that has a fairly large effect on the total sample size.

We also find that different versions of statistical packages, such as Stata 16 and Stata 17 in our case, can substantially affect outcomes, highlighting the importance of using the exact software version when testing computational reproducibility.

Overall, having considered the replication process and results, we conclude that this replication exercise supports the main conclusions of the original paper.

References

Chernozhukov, Victor, et al. (2018). "Double/debiased machine learning for treatment and structural parameters." *The Econometrics Journal*, 21(1): C1–C68.

Henry, Emeric, Ekaterina Zhuravskaya, and Sergei Guriev (2022). "Checking and sharing alt-facts." *American Economic Journal: Economic Policy*, 14(3): 55–86.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1): 267–288.

Table 1: Average treatment effects (intent)

		Replication variants				
	in *AEJ*	O	V	S	L	C
Panel A: Average treatment effect of sharing alt-facts						
Imposed Fact-Check	−0.050	−0.04999	−0.04999	−0.05025	−0.04011	−0.04499
	(0.017)	(0.01720)	(0.01720)	(0.01718)	(0.01492)	(0.01716)
Voluntary Fact-Check	−0.052	−0.05202	−0.05202	−0.05069	−0.04341	−0.04831
	(0.017)	(0.01743)	(0.01743)	(0.01741)	(0.01479)	(0.01646)
Observations	2,078	2,078	2,078	2,090	2,266	2,146
R-squared	0.166	0.16630	0.16630	0.17471		0.16941
Mean DV, alt-facts treatment	0.161	0.16087	0.16087	0.16259	0.14675	0.15302
p-value, imposed equal voluntary	0.899	0.89925	0.89925	0.97826	0.81112	0.83662
Panel B: Average treatment effect on sharing fact-checking						
Imposed Fact-Check	−0.039	−0.03896	−0.03896	−0.03642	−0.02510	−0.04284
	(0.018)	(0.01843)	(0.01843)	(0.01841)	(0.01624)	(0.01877)
Observations	1,388	1,388	1,388	1,395	1,484	1,401
R-squared	0.131	0.13128	0.13128	0.13722		0.12312
Mean DV imposed Fact-Check treatment	0.160	0.16017	0.16017	0.16092	0.14303	0.16717
Panel C: Actual compared to predicted sharing for viewers and non-viewers						
Imposed Fact-Check	−0.051	−0.05090	−0.05090	−0.05111	−0.04828	−0.04463
	(0.018)	(0.01762)	(0.01762)	(0.01761)	(0.01693)	(0.01717)
Voluntary Fact-Check: Viewer	−0.025	−0.02538	−0.02538	−0.02199	−0.02757	−0.02292
	(0.024)	(0.02445)	(0.02445)	(0.02445)	(0.02422)	(0.02359)
Voluntary Fact-Check: Non-Viewer	−0.072	−0.07196	−0.07196	−0.07230	−0.05786	−0.06591
	(0.019)	(0.01860)	(0.01860)	(0.01854)	(0.01785)	(0.01744)
Observations	2,018	2,018	2,018	2,030	2,118	2,146
R-squared	0.069	0.06878	0.06878	0.07174		0.06368
Mean DV Alt-Facts Treatment	0.0023	0.00229	0.00229	0.00238	−0.00139	0.00000
Mean predicted Alt-Facts Treatment	0.159	0.15857	0.15857	0.16021	0.15049	0.15302
Mean predicted imposed Fact-Check	0.159	0.15900	0.15900	0.16168	0.15803	0.16128
Mean predicted Voluntary T, Viewer	0.190	0.19030	0.19030	0.19441	0.18848	0.19186
Mean predicted Voluntary T, Nonviewer	0.144	0.14390	0.14390	0.14353	0.13645	0.13941

Note: This table replicates column 4 of Table 2, Henry et al. (2022). "DV" is short for "Dependent Variable." The replication variants are "O," which reproduces with the same code as HZG: "V" which uses the same code as "O" but is executed with Stata SE 17.0 instead of Stata SE 16.1 (used for the other variants); "S" removes the exclusion of speeders that used less than 250 seconds; "L" adopts the Chernozhukov et al. (2018) estimator (as implemented by the Stata "xporegress" command, default settings); and "C" regularize the set of pre-treatment variables in the two estimation steps. Robust standard errors in parentheses.

Table 2: Average treatment effects (action)

		Replication variants				
	in *AEJ*	O	V	S	L	C
Panel A: Average treatment effect of sharing alt-facts						
Imposed Fact-Check	−0.021	−0.02096	−0.02096	−0.02100	−0.01833	−0.02173
	(0.007)	(0.00699)	(0.00699)	(0.00694)	(0.00585)	(0.00668)
Voluntary Fact-Check	−0.026	−0.02618	−0.02618	−0.02612	−0.02219	−0.02415
	(0.006)	(0.00624)	(0.00624)	(0.00609)	(0.00535)	(0.00583)
Observations	2,078	2,078	2,078	2,090	2,266	2,146
R-squared	0.101	0.10075	0.10075	0.10031		0.10856
Mean DV alt-facts treatment	0.0509	0.05085	0.05085	0.05053	0.14675	0.04809
p-value, "imposed" equal "voluntary"	0.379	0.37927	0.37927	0.37675	0.44053	0.67153
Panel B: Average treatment effect on sharing fact-checking						
Voluntary Fact-Check	0.004	0.00368	0.00368	0.00468	0.00125	−0.00015
	(0.007)	(0.00653)	(0.00653)	(0.00639)	(0.00601)	(0.00685)
Observations	1,388	1,388	1,388	1,395	1,553	1,401
R-squared	0.073	0.07272	0.07272	0.08256		0.06923
Mean DV imposed Fact-Check treatment	0.0320	0.03197	0.03197	0.03183	0.03147	0.03444
Panel C: Actual compared to predicted sharing for viewers and non-viewers						
Imposed Fact-Check	−0.021	−0.02094	−0.02282	−0.02290	−0.02118	−0.02166
	(0.007)	(0.00699)	(0.00702)	(0.00697)	(0.00645)	(0.00668)
Voluntary Fact-Check: Viewer	−0.020	−0.02035	−0.02132	−0.02070	−0.02122	−0.01971
	(0.008)	(0.00818)	(0.00819)	(0.00791)	(0.00817)	(0.00781)
Voluntary Fact-Check: Non-Viewer	−0.030	−0.03048	−0.02994	−0.03032	−0.02687	−0.02722
	(0.007)	(0.00659)	(0.00657)	(0.00637)	(0.00626)	(0.00610)
Observations	2,078	2,078	2,018	2,030	2,211	2,146
R-squared	0.063	0.06899	0.07192	0.08007		0.07020
Mean DV Alt-Facts Treatment	0.0019	0.00187	0.00185	0.00059	−0.00007	0.00000
Mean predicted Alt-Facts Treatment	0.0490	0.04898	0.04900	0.04994	0.04729	0.04809
Mean predicted imposed Fact-Check	0.0480	0.04791	0.04977	0.05128	0.04675	0.04977
Mean predicted Voluntary T, Viewer	0.0592	0.05975	0.05995	0.06129	0.05942	0.05957
Mean predicted Voluntary T, Nonviewer	0.0442	0.04497	0.04547	0.04699	0.04362	0.04454

Note: This table replicates column 8 of Table 2, Henry et al. (2022). "DV" is short for "Dependent Variable." The replication variants are "O," which reproduces with the same code as HZG: "V" which uses the same code as "O" but is executed with Stata SE 17.0 instead of Stata SE 16.1 (which used for the other variants); "S" removes the exclusion of speeders that used less than 250 seconds; "L" adopts the Chernozhukov et al. (2018) estimator (as implemented by the Stata "xporegress" command, default settings); and "C" regularize the set of pre-treatment variables in the two estimation steps. Robust standard errors in parentheses.