

DISCUSSION PAPER SERIES

DP14738
(v. 3)

Checking and Sharing Alt-Facts

Emeric Henry, Ekaterina Zhuravskaya and Sergei
Guriev

PUBLIC ECONOMICS

CEPR

Checking and Sharing Alt-Facts

Emeric Henry, Ekaterina Zhuravskaya and Sergei Guriev

Discussion Paper DP14738
First Published 10 May 2020
This Revision 27 June 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Emeric Henry, Ekaterina Zhuravskaya and Sergei Guriev

Checking and Sharing Alt-Facts

Abstract

Using an online randomized experiment in the context of the 2019 European elections campaign in France, we study how fact-checking affects sharing of false news on social media. We exposed a random sample of French voting-age Facebook users to statements on the role of the European Union made by the far-right populist party Rassemblement National. A randomly selected subgroup of participants was also presented with fact-checking of these statements; another subgroup was offered a choice whether to view the fact-checking information. Then, all participants could choose whether to share the false statements on their Facebook pages. We show that (i) both imposed and voluntary fact-checking reduce sharing of false statements by about 45%; (ii) the size of the effect is similar between imposed and voluntary fact-checking; and (iii) each additional click required to share false statements substantially reduces sharing. These results inform the debate about policy proposals aimed at limiting propagation of false news on social media.

JEL Classification: D8, D91

Keywords: alternative facts, fake news, Sharing, social media, Facebook

Emeric Henry - emeric.henry@sciencespo.fr
Sciences Po Paris and CEPR

Ekaterina Zhuravskaya - ekaterina.zhuravskaya@psemail.eu
Paris School of Economics (EHESS) and CEPR

Sergei Guriev - sergei.guriev@sciencespo.fr
Sciences Po and CEPR

Acknowledgements

The order of co-authors' names is random: <https://www.aeaweb.org/journals/policies/random-authororder/search>; confirmation code PF8HX0v-pbSv. This experiment obtained the approval of the internal Institutional Review Board of the Paris School of Economics; the approval number is IRB 2019-005. The experiment was preregistered at the AEA RCT Registry; registry number AEARCTR-0004060. We thank Yoann Renard for excellent research assistance. Sergei Guriev thanks Sciences Po's Scientific Advisory Board for financial support. We are grateful to three anonymous referees for their excellent comments and suggestions. We also thank Ruben Enikolopov, Irena Grosfeld, Matthew Notowidigdo, Maria Petrova, David Stromberg, and participants at the CEPR Conference on Populism, RASA Annual Conference, and research seminars at Sciences Po and Paris School of Economics for helpful comments.

Checking and Sharing Alt-Facts*

Emeric Henry  Ekaterina Zhuravskaya  Sergei Guriev

June 27, 2021

Abstract

During the 2019 European elections campaign, we exposed a random sample of French voting-age Facebook users to false statements by a far-right populist party. A randomly selected subgroup was also presented with fact-checking of these statements; another subgroup was offered a choice whether to view the fact-checking. Participants could then share these statements on their Facebook pages. We show that (i) both imposed and voluntary fact-checking reduce sharing of false statements by about 45%; (ii) the size of the effect is similar between imposed and voluntary fact-checking; and (iii) each additional click required to share false statements substantially reduces sharing.

(JEL D8, D91)

(Keywords: Alternative facts, Fake news, Sharing, Fact-checking, Facebook)

*Henry: Department of Economics, Sciences Po, 28 rue des Saints Peres, 75007 Paris, France and CEPR (e-mail: emeric.henry@sciencespo.fr); Zhuravskaya: Paris School of Economics (EHESS) 48 boulevard Jourdan, 75014 Paris, France and CEPR (e-mail: ekaterina.zhuravskaya@psemail.eu); Guriev: Department of Economics, Sciences Po, 28 rue des Saints Peres, 75007 Paris, France and CEPR (e-mail: sergei.guriev@sciencespo.fr). The order of co-authors' names is random: <https://www.aeaweb.org/journals/policies/random-author-order/search>; confirmation code PF8HX0v-pbSv. This experiment obtained the approval of the internal Institutional Review Board of the Paris School of Economics; the approval number is IRB 2019-005. The experiment was preregistered at the AEA RCT Registry; registry number AEARCTR-0004060 (Guriev et al., 2019). We thank Yoann Renard for excellent research assistance. Sergei Guriev thanks Sciences Po's Scientific Advisory Board for financial support. We are grateful to three anonymous referees for their excellent comments and suggestions. We also thank Ruben Enikolopov, Irena Grosfeld, Matthew Notowidigdo, Maria Petrova, David Stromberg, and participants at the CEPR Conference on Populism, RASA Annual Conference, and research seminars at Sciences Po and Paris School of Economics for helpful comments.

In January 2017, when asked to comment on clearly false statements by Donald Trump’s press secretary, Sean Spicer, the president’s counselor Kellyanne Conway famously said that these were not falsehoods but “alternative facts.” Since then, the notion of “alternative facts” (or alt-facts, for short) took on a life of its own describing false or misleading statements by politicians. To limit the dissemination of alt-facts, some countries, such as Germany and France, have introduced laws allowing regulators to block social network accounts and sites. Many media and independent organizations have started large-scale fact-checking efforts (Politifacts.com in the United States; Agence France Presse, Libération, and Le Monde in France; Pagella Politica in Italy). These efforts face two major challenges. First, they may be too slow, as false news spread particularly fast (Vosoughi et al., 2018). Second, ex post fact-checking may be ineffective, since correcting beliefs after an audience’s exposure to false or misleading statements is very hard (Swire et al., 2017; Nyhan et al., 2019; Barrera et al., 2020). The key to limiting the impact of alt-facts is therefore to hinder their propagation on social media.

On May 26, 2020, Twitter took a step in this direction. When Trump tweeted that *“There is NO WAY (ZERO!) that Mail-In Ballots will be anything less than substantially fraudulent,”* Twitter added a fact-checking link at the bottom of Mr. Trump’s tweet; the link, in blue lettering and punctuated by an exclamation mark, referred to the CNN website where Trump’s claim was fact-checked and disproved (Conger and Alba, 2020). This was a conscious decision by Twitter to apply its recent policy to label tweets that would need fact-checking, and it was the first time it was used for such a high-profile politician, a user with 82 million followers (Conger, 2020a). The label did not state that the original tweet was false; it literally said *“Get the facts about mail-in voting.”* Moreover, it did not directly show users the evidence on mail-in voting—it simply provided them with an option to see it by clicking on the link. During the November 2020 election vote count, Twitter went further, flagging 11 of Trump’s tweets about election results as misleading. It hid the content of these tweets under a blue label *“Some or all of the content shared in this Tweet is disputed and might be misleading about an election or other civic process,”* adding a link to fact-checking information under *“Learn more”* and a link to the original tweets under *“View”* (Conger, 2020b).

How much does adding such fact-checking information to the original alt-fact messages affect users’ sharing of these alt-facts on social media? In this paper, we study this question using a randomized experiment involving actual sharing on Facebook, the largest social network. We address the following questions: Does providing fact-checking information reduce sharing of alt-facts, and thus, their propagation on social media? If so, is it as effective to let the individuals voluntarily access the fact-checking information, similarly to what Twitter chose to do, as to force fact-checking on them? What drives users’ choices to view the fact-checking and to share the alt-facts? Does adding clicks in order to share content online affect sharing, and if so, how?

In May 2019, in the context of the European Parliament elections, using the Qualtrics platform, we conducted an online survey experiment with 2,537 French voting-age Facebook users (for the data, see Henry et al., 2021). All the participants were shown two misleading

statements about the European Union (EU) made by the leaders of the far-right party Rassemblement National (RN), with links to the exact source.¹ The first claim stated that 87% of the French laws come from European directives; the second stated that the EU wants to attract 50 million immigrants to Europe by 2050. After seeing both of these statements, one-third of the subjects, randomly drawn, were exposed to fact-checking information compiled from media sources (we refer to this treatment group as Imposed Fact-Check) while another third was given the choice of accessing or not accessing this fact-checking information (Voluntary Fact-Check treatment group). The remaining third (Alt-Facts treatment group) was not shown fact-checking information and was not given an option to access it.

After being exposed to the alt-facts and (possibly) to the fact-checking, participants had the opportunity to share the alt-facts—the two statements by RN leadership—on their Facebook pages. First, they were asked within the survey whether they wanted to share this content on their Facebook page. Those who answered “yes” were taken to a screen with a Facebook “share” button. If they clicked on this button, they were brought to an external website that showed exactly how the information would appear on their Facebook page, where they were asked to reconfirm sharing by clicking again on the Facebook share button, which brought them to their Facebook pages.

The first question, which inquired about the intention to share alt-facts, was concrete—it specifically mentioned the Facebook page of the participants: “*Do you want to share (...) on your Facebook page?*” Therefore, the positive answer could be considered an expression of real willingness to share rather than an answer to a hypothetical question, which were used in other studies of determinants of sharing content online (e.g., Altay et al., 2020; Fazio, 2020; Pennycook et al., 2020a). Answering “yes” prompted the appearance of a Facebook share button. When participants clicked on that button, they did not know that it would bring them to another page, where they would need to reconfirm sharing, instead of bringing them directly to Facebook.² Thus, throughout the paper, we consider two outcomes: (i) a positive answer to the first question as the expression of real intent to share, which we refer to as the *sharing intention*, and (ii) clicking on the Facebook share button for the first time as the real action of sharing, which we refer to as the *sharing action*.

All 2,537 participants were exposed to the alt-facts, which they could share on Facebook. Among them, 302 participants expressed a sharing intention, i.e., answered “yes” to the question on whether they wanted to share this content on their Facebook page. Of these 302 participants, 82 took the sharing action, i.e., clicked on the Facebook share button, and only 21 chose to reconfirm for the second time, which resulted in posting the content on their Facebook pages. Overall, 11.9% of participants exposed to alt-facts expressed a sharing intention, 3.2% took

¹Until June 2018, Rassemblement National (the National Rally) was called Front National (the National Front). Marine Le Pen was the party’s president both before and after the renaming. We use the acronym RN to refer to the party throughout the rest of the paper.

²We introduced this additional (third) click in order to (a) limit the circulation of RN’s false narratives on Facebook as a result of our experiment and (b) study the effect of introducing a cost of an additional click on sharing decisions.

the sharing action, and about 0.8% reconfirmed sharing action for the second time. We find that, first, there is a large gap between self-reported sharing intention and the actual action of sharing, with much lower rates of clicking on the Facebook share button compared to the self-reported intention to share. Furthermore, the requirement to reconfirm the sharing action with a second click on the Facebook share button also substantially reduced sharing. Thus, the first result of our study is that even a very small cost associated with a need to reconfirm the original choice significantly and substantially reduces sharing. This finding provides a rationale for Twitter’s actions that hid Donald Trump’s tweets under a warning label as, to view the original tweet, users had to make an additional click (Conger, 2020a,b).

The central result of our paper comes from the comparison of sharing behavior among the three groups of participants: those who were exposed to alt-facts only, those who were exposed to alt-facts and fact-checking, and those who were given an option to view fact-checking after being exposed to alt-facts. We find that both voluntary and imposed fact-checking substantially and significantly reduce sharing of alt-facts. In our experiment, 14.7% in the Alt-Fact treatment group expressed an intent to share, but only 10.2% in the Imposed Fact-Check treatment group and only 10.8% in the Voluntary Fact-Check treatment group did so. Thus, fact-checking (imposed or voluntary) reduced intention to share alt-facts by more than 25%. The treatment effects for the sharing-action outcome are even stronger: 4.6% of subjects in the Alt-Facts treatment group clicked on the Facebook share button, but only 2.6% in the Imposed Fact-Check treatment group and only 2.5% in the Voluntary Fact-Check treatment group did so. Thus, actual sharing of alt-facts decreased by about 43% to 45% as a result of the imposed and voluntary fact-checking. For sharing intentions and sharing actions alike, the difference between the impacts of Imposed and Voluntary Fact-Check treatments is small and is not statistically significant.

Those participants who viewed the fact-checking information were also given an opportunity to share fact-checking information on Facebook. We show that the users shared the fact-checking information at similar rates regardless of whether fact-checking was imposed on them or was offered as an option: 3.15% of the Imposed Fact-Check participants and 3.07% of the Voluntary Fact-Check participants chose a sharing action for fact-checking information, i.e., they clicked on the Facebook share button to share fact-checking content. The difference between the actual sharing rates of fact-checking across the two treatments is small and statistically insignificant. (The respective sharing-intention rates for fact-checking are 14.3% in the Imposed Fact-Check treatment group and 11.5% in the Voluntary Fact-Check treatment group.)

We conclude that giving users the option to view fact-checking information is as effective as imposing fact-checking on them in both limiting the propagation of false news and spreading fact-checking information. This is important from a policy perspective, since in most real-world settings it is much easier to offer social media users an option to view fact-checking information than it is to force it on them.

At first glance, the finding that the voluntary and imposed fact-checking treatments have

the same impact on sharing alt-facts may seem puzzling. The effect of the Voluntary Fact-Check treatment is a weighted average of the effect on those who chose to view the fact-checking information (“Viewers,” who constitute 39% of the Voluntary Fact-Check group) and the effect on those who did not (“Nonviewers”). One could expect the effect on Viewers to be similar to the effect on the Imposed Fact-Check participants, as both got exposed to fact-checking information; one could also expect no effect on Nonviewers (who, after all, did not see the fact-checking information). This is not what we find: on average, Viewers of fact-checking information are more likely to share alt-facts than Nonviewers.

The key to the puzzle is the endogeneity of the choice to view the fact-checking information. The self-selection into Viewers vs. Nonviewers is correlated with the ex ante propensity to share alt-facts. To demonstrate this, we use the adaptive LASSO method (Zou, 2006) on the full battery of pretreatment characteristics and the actual sharing behavior in the Alt-Facts treatment group to predict the ex ante propensity to share alt-facts for the Imposed Fact-Check and Voluntary Fact-Check participants.³ We find that Viewers have a significantly higher ex ante propensity to share alt-facts than Nonviewers. Then, we compare the actual sharing decisions with the predicted ex ante propensity to do so across different groups of participants; we find that for both Viewers and Nonviewers there is a significant decrease in actual sharing compared to the ex ante propensity to share. Among Viewers, the rate of actual sharing is 39% lower than their predicted propensity to share; among Nonviewers, the respective number is 60.2%. Thus, even though Nonviewers decided not to view the fact-checking, the very knowledge of the existence of fact-checking greatly reduced their sharing of alt-facts. The fact that both Viewers and Nonviewers reduced their sharing of alt-facts sheds light on why the Imposed Fact-Check and Voluntary Fact-Check treatments yield similar results. This is also important for policy implications, because in a real-world setting social media platforms cannot control who chooses to view fact-checking when doing so is voluntary.

Those who chose to view fact checking in the Voluntary Fact-Check treatment group not only have a higher ex ante propensity to share alt-facts but also are more likely ex ante to share fact-checking information.⁴ Ex post—after exercising their option to view the fact-checking—Viewers share fact-checking at a substantially higher rate compared both to how much they share alt-facts and to their predicted propensity to share fact-checking.

Our findings have important policy implications, which we discuss in detail after we

³The following factors predict the sharing action. Men shared alt-facts more than women. Those reporting that they share on social media to influence others shared alt-facts more often than those who refer to other motives of sharing. People who report to be religious and to have donated blood also shared alt-facts more. The ideological alignment with alt-facts also plays a key role: those who voted for Le Pen in the last presidential election were more likely to share alt-facts, while those who feel that their are closer to the European government than to the French government were less likely to share alt-facts.

⁴As in the case of ex ante propensity to share alt-facts, we predict ex ante propensity to share fact-checking using pretreatment characteristics and the actual sharing behavior in the Imposed Fact-Check treatment group. Some of the predictors of sharing fact-checking are similar to those of sharing alt-facts: for example, being male and naming influence as a usual reason to share information on social media are important predictors of sharing both alt-facts and fact-checking. Some predictors, however, are different: less religious people, people who have a positive image of the EU, and people who state that reciprocity is a usual reason for them to share are more likely to share fact-checking information if exposed to it.

present the results. Propagation of false narratives on social media affects political beliefs and voting choices and, therefore, has real consequences for socioeconomic outcomes. We show that providing (voluntary) access to fact-checking substantially reduces sharing of alt-facts. As with any costly information, fact-checking is a public good: individual returns to fact-checking are much lower than the social returns. Therefore, our findings imply the need for supporting fact-checking and requiring the platform to accompany influential alt-facts posts with a link to fact-checking. Our results also suggest that asking users to reconfirm their willingness to share alt-facts as well as non-fact-checked content may substantially reduce dissemination of alt-facts.

Our work is related to the literature on circulation of false news on social media. [Allcott and Gentzkow \(2017\)](#) show that fake stories were intensely shared on Facebook during the 2016 U.S. presidential election campaign. [Vosoughi et al. \(2018\)](#) show that false news spread faster than real news. [Mosleh et al. \(2020\)](#) show that sharing of news on Twitter is negatively correlated with the veracity of the news headline. [Guess et al. \(2019\)](#) examine what factors correlate with the sharing of alt-facts, highlighting the role of political alignment with the news and the effect of old age. (We also find that political alignment matters, but do not find a strong effect of age.) [Pennycook et al. \(2019\)](#) and [Pennycook et al. \(2020b\)](#) show that even though individuals are good at determining the accuracy of news, and in particular can distinguish fake news from real news, accuracy is not the key factor for their self-reported intention to share; the authors posit that even minimal interventions can make accuracy more salient. [Fazio \(2020\)](#) shows that asking participants to explain whether a headline is true or false reduces their sharing intention of false news. [Altay et al. \(2020\)](#) show that, in addition to accuracy, interest in the news also plays a role in self-reported sharing intentions.

We also contribute in four ways to a small, but growing literature that examines the impact of fact-checking on sharing false news. [Yaqub et al. \(2020\)](#) show that putting labels below a news headline—indicating that the news has been fact-checked and shown to be false (even though the actual fact-check is not shown)—decreases the self-reported intention to share. [Pennycook et al. \(2020a\)](#) carried out an online experiment where the participants were shown true and false statements. They find that adding the “false” label to a statement significantly reduces participants’ self-reported intention to share the statement on social media. So first, unlike these papers, which examine the answers to a purely hypothetical question, such as “*We are interested in whether you would consider sharing these stories on social media,*” we study actual sharing on Facebook. Our two main outcomes are the answer to a direct, concrete question that refers to a specific sharing channel: “*Do you want to share [...] on your Facebook page?*” and the subsequent actual clicking on the Facebook share button.⁵ Second, we compare the effect of imposed fact-checking to voluntary fact-checking on the sharing of alt-facts, a relevant policy issue that has not been previously studied. Along these lines, our paper is

⁵In Online Appendix B, we provide precise formulations of questions on willingness to share in all these studies of sharing false news and discuss the differences with our measures. The English translation of our questions is presented in Figure 1. In the Online Appendix B, we also provide the original formulation of our questions in French.

the first to analyze the strategic choice to access fact-checking if users are given an option to fact-check.⁶ Third, we also consider sharing fact-checking information as an outcome; we show that access to fact-checking leads to its dissemination on social media. Finally, we highlight the policy implications of our results on the impact of voluntary fact-checking and requiring users to make additional clicks to confirm their intention to share.

Taken together, the results of our randomized experiment with real Facebook sharing decisions deliver an optimistic message. Although the previous literature has shown that fact-checking cannot undo initial impressions left by false statements [Barrera et al. \(2020\)](#); [Nyhan et al. \(2019\)](#); [Swire et al. \(2017\)](#), our paper highlights an important role that fact-checking plays in limiting propagation of false news.

The rest of the paper is organized as follows. Section I describes the design of the experiment. Section II presents the main results on how fact-checking affects sharing of alt-facts and provides evidence on the link between the decisions to fact-check and to share content. Section III discusses alternative interpretations of the effect of exposure to fact-checking on sharing of alt-facts. Section IV discusses the policy implications. Section V concludes.

I Experimental design

Using the context of the European Parliament elections held in France on May 26, 2019, we focus on the misleading statements made by leading members of the far-right RN party. The objective of these elections was to select France’s 79 members of the European Parliament. The elections were held in all EU countries but followed different rules in each country. In France, voters could choose between different lists of ranked candidates. Proportional to their vote share, seats were allocated to lists that obtained at least 5% of the vote.

This election was essentially a contest between the EU-skeptic RN party (Marine Le Pen’s party) and the pro-EU La République En Marche (LREM, Emmanuel Macron’s party). It was a rematch of the second round of the 2017 presidential election, in which Macron beat Le Pen. In 2019, RN obtained 23% of the votes and 23 seats in the European Parliament, while LREM obtained 22% and 23 seats.

I.A Alt-Facts and Fact-Checking Used in the Experiment

In our experiment, we use two alt-facts statements, each related to a central theme of RN’s anti-EU campaign. In the first statement, Le Pen argues that the EU promotes immigration, which the RN leadership believes to be harmful for France. Specifically, in a televised interview on France Info on June 25, 2018, Le Pen said: *“The European Union wants immigration. It*

⁶This relates to a growing theoretical literature on strategic sharing of information in networks. [Bloch et al. \(2018\)](#) study the circulation of rumors in a social network, with unbiased agents evaluating their beliefs about veracity of a news and potentially blocking them. See also [Campbell et al. \(2019\)](#). [Bénabou et al. \(2019\)](#) study the circulation of arguments justifying certain actions. [Abreu and Jeon \(2019\)](#) explore how media strategically positioning their news to induce sharing.

has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 million immigrants by 2050.”⁷ Participants were asked to read this statement; they could also watch the original video of the interview. This statement was fact-checked by *Libération*, a major French newspaper: it reported that the figure “50 million” first appeared in a June 8, 2018, statement by the European Commissioner for Migration, Home Affairs and Citizenship, Dimitris Avramopoulos, who said, “*We know Europe is an ageing continent. Without migration, the active population will decline by more than 20 million in the fifteen upcoming years. By 2060, the active population would have decreased by 50 million.*” The newspaper noted that this observation does not in any way imply that these population losses need to be compensated by immigrants. Moreover, the European Commission does not have the authority to make decisions on issues related to immigration.

The second statement was made by Jordan Bardella, the politician ranked first on the RN list for the 2019 European election. In an interview on Sud Radio on June 4, 2018, Bardella said, “*We have to regain control of our budgets, regain control of our laws. I remind you that 87% of our laws, laws that are voted, come from European directives.*”⁸ Participants were asked to read this text. Bardella’s statement was fact-checked multiple times, including by France TV Info and Public Senat. Töller (2010) states that “most of the studies have shown a relatively low share of laws coming from European legislation, 15.5% in United Kingdom, 14% in Denmark, 10.6% for Austria, between 3 and 27% for France.” All studies cited in Töller (2010) report numbers significantly below 50%, well below the 87% figure quoted by Bardella.

In Online Appendix C, we provide the text of the English translation of the alt-facts treatment and the imposed and voluntary fact-checking treatments that were presented to the participants.⁹

I.B The Experiment Setup

We ran our experiment in the week preceding the European elections using the Qualtrics online platform. A total of 2,537 French voting-age Facebook users were contacted by Qualtrics via email.

We stratified our sample on education and gender by treatment. The sampling quotas were designed to make the sample roughly representative of the French adult population eligible to vote: we imposed quotas on gender (50% male, 50% female) and on education (high school and below 72%, undergraduate degree 12%, graduate degree 16%).

At the start of the survey, participants were given a brief introduction indicating the

⁷The original statement in French was, “*L’Union européenne souhaite de l’immigration. Elle l’a dit à plusieurs reprises par l’intermédiaire de la bouche, d’ailleurs, de beaucoup de ses commissaires européens. Ils sont même allés jusqu’à dire 50 millions d’immigrés d’ici 2050.*”

⁸The original statement in French was, “*Il nous faut retrouver la maîtrise de nos budgets, retrouver la maîtrise de nos lois. Je rappelle que 87% de nos lois, des lois qui sont votées, viennent de directives européennes.*”

⁹We chose to limit ourselves to two statements only (as opposed to Pennycook et al., 2020a; Yaqub et al., 2020) since, in addition to these statements, a subset of participants was exposed to extensive fact-checking articles, making the overall treatment quite lengthy.

survey’s focus on social-media behavior. It also stated that only aggregate results would be published. There was no mention of any political party or political candidate. The introductory page allowed participants to drop out at this stage. The institutional affiliations of the authors were not specified, to avoid the appearance of potential ideological biases of the experiment designers.

The survey consisted of three parts: (1) We asked all participants a series of questions regarding their socioeconomic characteristics, such as age, gender, education, income, and religion. We then asked participants about their use of social media and, in particular, their motivations for sharing content on Facebook. We also asked questions to measure their level of altruism, reciprocity, and image concerns, social preferences that can affect sharing decisions. Finally, we asked them about voting behavior in the 2017 presidential elections as well as a question measuring their feelings towards the European Union. (2) We randomly assigned participants to informational treatments, which we describe in the next section. (3) To measure voting intentions and attitudes toward the European Union, we also included a factual question on participants’ beliefs regarding the share of French laws coming from European directives. The translation of the text of each treatment into English is presented in Online Appendix C. The full set of pre- and posttreatment questions is reported in Online Appendix D.

I.C Treatments

Participants were randomly allocated to three groups. Each group was exposed to a different informational treatment:

1. **Alt-Facts treatment.** Participants were exposed to alt-facts (the two false statements by RN presented above). After seeing the alt-facts, they were asked whether they wanted to share them on their Facebook page, and they were given the opportunity to actually do so (details are provided below). Then, they were asked whether they wanted to anonymously share the content with 100 other individuals.¹⁰
2. **Imposed Fact-Check treatment.** Participants were first exposed to the alt-facts, then they were exposed to the fact-checking of these two RN statements. After having seen both the alt-facts and the fact-checking, participants were asked whether they wanted to share the alt-facts on their Facebook page, and they were given the opportunity to actually do so. Then, as in the Alt-Facts treatment, they were asked whether they wanted to anonymously share the content with 100 individuals. Finally, they were asked whether they wanted to share the fact-checking information on Facebook, and again they were given the opportunity to actually do so.
3. **Voluntary Fact-Check treatment.** Participants were first exposed to alt-facts. They were then offered a choice to view or not to view the fact-checking of the alt-facts. After

¹⁰Those individuals were the participants in a follow-up survey-experiment that did take place, but as it has a different focus, we do not explore it in this paper.

they made their choice, those who chose not to view the fact-checking (“Nonviewers”) proceeded as participants of the Alt-Facts Treatment; and those who chose to view fact-checking (“Viewers”) proceeded as participants of the Imposed Fact-Check Treatment group. In other words, all participants of the Voluntary Fact-Check Treatment group were given the opportunity to share the alt-facts on their Facebook page and were also asked if they wanted to share the alt-facts with 100 participants taking the survey after them. “Viewers” were in addition given the opportunity to share the fact-checking information on their Facebook page.

The text of each treatment is presented in Online Appendix C.

I.D Sharing Procedure and the Measures of Sharing on Facebook

The procedure of sharing on Facebook was as follows:

- The participants were first asked within the Qualtrics survey whether they wanted to share on Facebook. Specifically, the question was: “Do you want to share this article ‘RN: the European Union controls 87% of our laws and wants to open the door to 50 million of immigrants’ containing the declarations made by Marine Le Pen and Jordan Bardella on your Facebook page?”¹¹ (We provide the image of how this question appeared on participants’ screens in English in Figure 1 and the screenshot of the original question in French in Online Appendix B.) Participants could answer “yes,” “no,” or “do not want to answer.” We use the answer to this question as our first outcome variable “Intent to share alt-facts on Facebook,” or *sharing intention*.¹²
- If participants answered “yes” to this first question, they were brought to a second window within the Qualtrics environment with a button that looked exactly like a Facebook “share” button. (See the image of how this question appeared on participants’ screens in English also in Figure 1 and the screenshot of the original question in French in Online Appendix B.) This Qualtrics page also stated that the users could log into Facebook if they were not already logged in. If they clicked on the Facebook share button, they were taken to a webpage (outside the Qualtrics environment) that showed the information to be shared—exactly as it would appear on Facebook (see Online Appendix C). Clicking on the Facebook share button within Qualtrics (i.e., for the first time) could not be directly recorded due to a technical constraint specific to the Qualtrics survey environment, but

¹¹In all our treatments, the two alternative facts were combined into one “article.” The participants could share either both alternative facts or none. The same was the case for fact-checking: the participants could only share fact-checking of both alt-facts or none at all. In reality, Facebook users can choose sharing any part of alt-facts or fact-checking articles.

¹²The typical formulation of the question on intention to share in related studies is much more hypothetical and does not refer to a specific platform. In Online Appendix B, we provide the formulations of the question for all the relevant studies cited in our paper. For instance, in Pennycook et al. (2020a), the question is, “We are interested in whether you would consider sharing these stories on social media (such as Facebook or Twitter).” We use the verb “want,” and we specifically refer to the participant’s Facebook page.

we could track the visits to the webpage within each hour using Google Analytics. We matched the Google Analytics information with the survey (see details of the procedure in Online Appendix E), which allows us to calculate the probability that each survey participant actually visited the webpage. We recorded this probability as “Action of sharing alt-facts on Facebook,” or *sharing action*; we use it as our second outcome variable. This variable equals the number of visits to the webpage within an hour divided by the number of “yes” answers to the “want to share” question within the same hour. Thus, there is a measurement error in the “sharing action” variable when it is not strictly zero or one. (In the left-hand-side chart of Online Appendix Figure A1, we present the distribution of this variable for respondents with positive intention to share alt-facts). However, as we use this variable only as the outcome variable (rather than the explanatory variable), this measurement error does not bias the estimated coefficients and only increases standard errors of the estimated effects, making it harder to detect significance. (As we show below, despite this measurement error, our estimates of the treatment effects are statistically significant.)

- On the webpage, the participant could click on the Facebook share button again to reconfirm sharing. Clicking brought them directly to Facebook. Based on these clicks, we have created the variable “Reconfirm sharing alt-facts on Facebook.” We could track the number of total shares of each URL on Facebook via share buttons with the count option activated on another page inaccessible to the participants. We could separately identify the shares by our survey participants and the reshares by the friends of these participants, following a procedure described in Online Appendix E.
- As mentioned above, participants who viewed the fact-checking information (in the Imposed Fact-Check or Voluntary Fact-Check treatment groups) could also share the fact-checking information on Facebook. The procedure was exactly the same as for sharing alt-facts. Therefore, for sharing fact-checking information, we also have three variables: “Intent to share fact-check on Facebook,” “Action of sharing fact-check on Facebook,” (see its distribution on the right-hand-side chart of Online Appendix Figure A1), and “Reconfirm sharing fact-check on Facebook,” which is the second click on the Facebook share button.

Our study involved actual sharing on real, personal Facebook pages. We were granted an IRB approval for three reasons: (1) we chose real public statements made by politicians rather than making up false content; (2) we committed to limiting the circulation of false narratives to a minimum, which led us to introduce additional steps such as asking the respondents to reconfirm the sharing decision; (3) we also committed to taking down the alt-facts webpage if sharing became excessive. After three days, we thus changed the content of the shared webpage, replacing the alt-facts with the corresponding fact-checking information. After five days, we deactivated the website. The overall number of shares and reshares of alt-facts in our

experiment (less than 800) is minuscule relative to the size of the audience that was exposed to the original alt-fact statements via TV and radio.

I.E Descriptive Statistics and the Balance Across Treatments

Columns 1 to 6 of Table 1 present the summary statistics for all the main pretreatment characteristics of respondents by treatment group. The last two columns present the p-values for the test of equality of means between the Alt-Facts treatment group and the two Fact-Check treatment groups. (We provide detailed descriptions of all the main variables in the Online Appendix E.) In the Online Appendix Table A1, we report the results of an omnibus test of randomization quality. We regress the treatment status with three outcomes: Alt Facts, Imposed Fact Check, Voluntary Fact Check, in a multinomial logistic regression, on all pretreatment characteristics and test for their joint significance. For the comparison between Alt Facts and Imposed Fact Check treatments, the p-value of this test is 0.74; between Alt Facts and Imposed Voluntary Fact Check, the p-value is 0.17; and for predicting any treatment status, the p-value is 0.26. Overall, randomization worked relatively well, even though there are some imbalances across treatment groups. Below, we show that none of the results on treatment effects depend on the inclusion or exclusion of any of the pretreatment characteristics from the set of covariates.

Our sample consists of only Facebook users, but it was not designed to be representative of the Facebook users in France (it was stratified to match the French adult population as a whole on education and gender). As shown in Table A2 in the Online Appendix, on a number of socioeconomic variables and measures of Facebook usage, the means in our sample are somewhat different from the average statistics for the population of French Facebook users. In 2019, Facebook had 35 million Facebook users in France, including 27 million daily users.¹³ In our sample, 70% are heavy Facebook users (they check Facebook several times a day) and 50% share often. On average, the participants in our experiment had 214 Facebook friends, whereas the average for Facebook users worldwide is 338. On average, our participants are older and more likely to be married than a mean French Facebook user; they are also less likely to have tertiary education.

II Results

II.A Aggregate Sharing and the Effect of an Additional Click

Before presenting our main results on the effects of the treatments, we provide aggregate statistics on sharing that demonstrate that the content we showed to our participants did make an impact. In addition to being shared by the participants, the alt-facts were also reshared by Facebook users outside of the experiment. Overall, alt-facts posts of our experiment participants

¹³See, for instance, an interview with Laurent Solly, a vice president at Facebook, at <https://www.lejdd.fr/Medias/Internet/il-y-a-plus-de-35-millions-dutilisateurs-de-facebook-en-france-3850928>, accessed on January 4, 2021.

were shared and reshared on Facebook 749 times in five days.¹⁴ The fact-checking was shared and reshared 52 times, reflecting the fact that substantially fewer people had the opportunity to share it (only 846 in the Imposed Fact-Check treatment group and 330 of those who chose to view fact-checking in the Voluntary Fact-Check treatment group). Still, the number of shares-cum-reshares per exposed experiment participant is much higher for alt-facts ($749/2,537=0.30$) than for the fact-checking ($52/1,176=0.04$). This is consistent with the literature, which finds that false news disseminates on social media much faster than true stories (e.g., Vosoughi et al., 2018). Figure 2 shows the cumulative sharing and resharing, from the time when the median respondent participated in the survey experiment.

Of the 2,537 participants in our experiment, 302 expressed a sharing intention, i.e., said that they want to share alt-facts on their Facebook page. Of those, 82 took the sharing action, i.e., clicked on the Facebook share button to share alt-facts and, therefore, visited the external website; of these 82, 21 ended up sharing on Facebook. Thus, at each step, the number of users continuing to the next step decreased dramatically. As shown in Panel A of Figure 3, this proportional decrease at each step is very similar across treatments. This demonstrates that, first, self-reported sharing intentions grossly overestimate actual sharing and, second, that even a very small cost (such as an additional click) can substantially reduce sharing. Panel B of the figure shows a similar pattern for the effect of introducing additional clicks for sharing the fact-checking information. Of 1,176 participants who viewed fact-checking information, 218 expressed the intent to share it, 53 clicked on the Facebook share button, and only 7 reconfirmed sharing on the external website by clicking on the share button for the second time. This suggests that the effect of an extra click is not related to the content of the information that users want to share.

The fact that only about 3% ($82/2,537=3.2\%$) of participants took the sharing action and less than 1% ($21/2,537=0.8\%$) of participants completed all the additional clicks to ultimately share the content on their Facebook page is an important result. The initial sharing of alt-facts is rare, especially if sharing requires additional clicks.¹⁵ On the other hand, Figure 2 shows that even if a very small number of users share alt-facts, their propagation on social media may take off quite rapidly. In Section IV we discuss policy measures that can limit posting and sharing of alt-facts on social media.

¹⁴Three days after the start of the experiment, we changed the content of the webpage, replacing the alt-facts with the fact-checking information; by that time, the posts with the link to this website had been shared 612 times. Even after that, more than 100 shares took place. This shows that many Facebook users share the posts with embedded links only on the basis of a headline without actually having clicked on the content of these links. This provides a rationale for Twitter’s recent “read before you retweet” experiment, which prevented a subset of its users from retweeting based on a headline without viewing the content of the article referred to in the tweet (see, for instance, Montti (2020) and also coverage of this attempt in the media: <https://www.theverge.com/2020/9/25/21455635/twitter-read-before-you-tweet-article-prompt-rolling-out-globally-soon>, accessed on January 4, 2021).

¹⁵Our results are specific to the alternative facts we expose our survey participants to but the magnitudes are comparable to those found in a nonexperimental study by Guess et al. (2019), who show that during the 2016 U.S. presidential election campaign, 5.2% of users shared at least one fake-news story and only 3.2% of users shared two or more fake-news stories.

II.B The Effect of Fact-Checking on Sharing Alt-Facts

Our first main result is that exposure to fact-checking substantially reduces sharing of alt-facts on social media. Our second main result is that the impact of fact-checking is the same whether viewing it is imposed or optional. To demonstrate this, we compare the means of our outcome variables across treatments. Panel A of Figure 4 presents the unconditional means for the sharing intention (on the left) and the sharing action (on the right) by treatment. We find that while 14.7% of subjects in the Alt-Facts treatment group expressed an intention to share alt-facts on their Facebook page, the mean rate of the sharing intention is only about 10.2% in the Imposed Fact-Check treatment group and only 10.8% in the Voluntary Fact-Check treatment group. The actual sharing of alt-facts also declines when fact-checking is available: 4.6% of subjects in the Alt-Fact treatment group took the sharing action, whereas only 2.6% in the Imposed Fact-Check treatment group and 2.5% in the Voluntary Fact-Check treatment group did so. Therefore, exposure to fact-checking (imposed or voluntary) reduces the intention to share alt-facts by 27% to 31% and the sharing action by 43% to 45%.¹⁶

Panel A of Table 2 presents these results in the form of OLS regressions. (Fuller regression output is presented in Online Appendix Table A3). The sharing outcomes are regressed on the dummies for the Imposed Fact-Check and Voluntary Fact-Check groups, leaving the Alt-Facts participants as the comparison group. The first four columns focus on the sharing intention as the outcome variable, and the last four columns focus on the sharing action. The regression results yield three takeaways. First, the effects of the Imposed Fact-Check and Voluntary Fact-Check treatments on both the sharing intention and the sharing action are statistically significant. Second, the table demonstrates that the average treatment effects of the Imposed Fact-Check and Voluntary Fact-Check treatments are indistinguishable from each other, as illustrated by the p-values for the test of the equality of the coefficients on the dummies for the two treatments, reported at the bottom of the table. Third, the comparison of the coefficients on the treatment dummies presented in different columns for each of the two outcomes illustrates that the choice of controls is essentially irrelevant, as one should expect due to randomization. In Columns 1 and 5, we include no controls. In Columns 2 and 6, we include only strata controls. In Columns 3 and 7, we add a number of demographic characteristics and self-reported (pretreatment) frequency of Facebook use, frequency of sharing on Facebook, and the log number of Facebook friends. In Columns 4 and 8, we control for *all* pretreatment characteristics, including the ones with a significant difference across treatments. The point estimates do not depend on the choice of controls.

The magnitudes of the treatment effects in our experiment are comparable to those identified in recent studies, which—in contrast to our study—focus solely on self-reported intention to share. Pennycook et al. (2020a) find that making participants rate the accuracy of a head-

¹⁶Despite the fact that the self-reported sharing intention is much larger than the actual sharing of alt-facts, we show that treatment effects do go in the same direction for the self-reported and actual sharing. This result is consistent with Mosleh et al. (2020), who show that self-reported intention to share content on social media is correlated with actual sharing.

line reduces the intention to share by 19%. [Yaqub et al. \(2020\)](#) show that putting an indicator flagging that the news was fact-checked and found to be false, reduces willingness to share by 43%. [Fazio \(2020\)](#) finds that pausing to consider whether a headline is true or false reduces the sharing intention of false news by 32%.¹⁷

We conclude that fact-checking significantly and substantially decreases sharing of alt-facts on Facebook. This effect is similar regardless of whether fact-checking is imposed on users or offered to them as an option. Since the average treatment effect of the Voluntary Fact-Check treatment is a weighted average of the sharing behavior of those who chose to exercise the option to view fact-checking and those who did not, we analyze the differences in their behaviors in Section II.D.

II.C The Effect of Imposed vs. Voluntary Fact-Checking on Sharing Fact-Checking Information

Exposure to fact-checking not only reduces sharing of alt-facts but also can result in sharing of fact-checking information itself. Panel B of Figure 4 illustrates the mean sharing of fact-checking information in the Imposed Fact-Check and Voluntary Fact-Check treatments. As in Panel A, the left graph presents the mean rates of the sharing intention, and the right graph presents those of the sharing action. We find that 14.3% of the subjects in the Imposed Fact-Check treatment group expressed an intention to share the fact-checking information; in the Voluntary Fact-Check treatment group, this number is 11.5%. Even though the mean rates of sharing intentions are different across the two treatments, the mean rates of sharing action are almost exactly the same: 3.15% in the Imposed Fact-Check treatment group and 3.07% in the Voluntary Fact-Check treatment group.

Panel B of Table 2 presents the corresponding regression results. (Fuller regression output is presented in Online Appendix Table A4.) We use the subsample of participants who could potentially share the fact-checking information, i.e., subjects in the Imposed Fact-Check and Voluntary Fact-Check treatment groups. We regress the sharing intention and sharing action on the dummy for the Voluntary Fact-Check treatment. The results confirm that, even though there is a significant difference in the intentions to share fact-checking information between the Imposed Fact-Check and Voluntary Fact-Check treatment groups, the rate of actual sharing of the fact-checking information is the same between these two treatment groups. This finding is surprising, as only 39% of participants in the Voluntary Fact-Check treatment group chose to view the fact-checking and thus had an option to share it. (In the next section, we show that this result is driven by the relationship between viewing and sharing decisions.)

Overall, in terms of sharing alt-facts and sharing fact-checking, the average behavior in the Imposed Fact-Check and the Voluntary Fact-Check treatment groups is remarkably similar.

¹⁷19% in [Pennycook et al. \(2020a\)](#) comes from Table S3, p. 26 of the Supplementary Materials: 0.0545/0.285. 32% in [Fazio \(2020\)](#) comes from the paper’s Finding 1 showing a reduction in the number of participants saying that they would be “likely,” “somewhat likely,” or “extremely likely” to share from 0.57 to 0.39.

II.D Viewing Fact-Checking Information and Sharing Behavior

The Voluntary Fact-Check treatment offered participants a choice to view the fact-checking information or not. Of the 846 participants in this treatment group, 330 (39%) exercised this option. In this section, we explore the difference in sharing behavior between Viewers and Nonviewers in the Voluntary Fact-Check treatment group.

In the charts on the left-hand side of each of the four panels of Figure 5, we summarize outcomes by treatment, but—in contrast to Figure 4, which presents the aggregate figures for the Voluntary Fact-Check treatment—in Figure 5, we present outcomes separately for Viewers and Nonviewers (the third and fourth bars in each chart, respectively). In the top row of the figure (Panels A.1 and A.2), we consider the sharing of alt-facts. As the left-hand-side charts of both A.1 and A.2 indicate, Viewers share alt-facts more than Nonviewers: both the rate of the intent to share alt-facts on Facebook (Panel A.1, left chart) and the rate of actual sharing (Panel A.2, left chart) are higher among Viewers than among Nonviewers. This may seem counterintuitive: one could expect that Viewers have a better understanding that the alt-facts are false compared to Nonviewers because they have actually read the fact-checking information.¹⁸ However, this hypothesis does not take into account the endogenous selection into viewing, which may be correlated with the individual’s propensity to share.

Therefore, we examine how the decision to view is related to the propensity to share alt-facts on Facebook ex ante, i.e., before the exposure to treatment. We predict the ex ante propensity to share alt-facts using all pretreatment characteristics and the decision to share alt-facts using the sample of participants in the Alt-Facts treatment group (who by design were neither exposed to fact-checking nor given the option of viewing fact-checking). Using the adaptive LASSO (Zou, 2006), we select the most important determinants of sharing alt-facts among pretreatment characteristics and we make an out-of-sample prediction about the ex ante propensity to share for each experiment participant. Columns 1 and 2 of Online Appendix Table A5 illustrate the main determinants of sharing alt-facts by presenting the results of OLS regressions, in which the sharing intention (Column 1) and the actual sharing of alt-facts (Column 2) in the sample of Alt-Facts treatment group are regressed on the variables selected by the adaptive LASSO.

The following factors are the most important in predicting the sharing of alt-facts: Those who voted for Le Pen in the second round of the 2017 presidential election are 20 percentage points more likely to declare an intention to share alt-facts and 7.3 percentage points more likely to take the sharing action, demonstrating the importance of ideological alignment on sharing. Similarly, those who report feeling closer to the European government than to other levels of government (and most notably to the national government) are 8 percentage points less likely to express the intent to share alt-facts about the loss of French sovereignty to the European Union and 3.9 percentage points less likely to actually share these alt-facts. Respondents who report that the opportunity to influence others motivates them to share content on Facebook are also

¹⁸In the Pre-Analysis Plan, we hypothesized a lower rate of sharing among Viewers.

significantly more likely to express a sharing intention and to take the sharing action: a one-standard-deviation increase in the importance of this motive to share increases the probability they will report an intention to share alt-facts by 6.3 percentage points and the probability they will take the sharing action by 1.3 percentage points. Religious individuals are 14.5 percentage points more likely than nonreligious individuals to report a sharing intention and 5.6 percentage points more likely to take the sharing action.¹⁹ Men, on average, intend to share alt-facts 5.3 percentage points more than women and actually share alt-facts 2.7 percentage points more. Finally, those who reported to have donated blood are 7.6 percentage points more likely to report a sharing intention and 2.6 percentage points more likely to take the sharing action.

In the charts on the right-hand side of Panels A.1 and A.2 of Figure 5, we summarize the resulting predicted values of the ex ante propensity to express intent to share alt-facts and to click on the share button, keeping the same scale for the vertical axis as on the left-hand-side charts. The mean of each predicted outcome is, by construction, exactly equal to the mean of the actual outcome in the Alt-Facts treatment group.²⁰ The predicted propensity to share alt-facts in the Imposed Fact-Check treatment group is also very close to—and statistically indistinguishable from—that in the Alt-Facts treatment group. This is what one should expect due to randomization and the balance across treatments in pretreatment characteristics. In contrast, there is a major difference between Viewers and Nonviewers in predicted propensity to share. Viewers have significantly higher ex ante propensity to share compared to Nonviewers: 18.8% vs. 13.6% for sharing intention, and 5.9% vs. 4.3% for sharing action, respectively. These differences are statistically significant, implying that those who decide to view fact-checking have pretreatment characteristics that predict substantially higher ex ante propensity to share. We illustrate this in Online Appendix Figure A2, which presents the means of the main pretreatment determinants of sharing of alt-facts separately for Viewers and Nonviewers in the Voluntary Fact-Check treatment group.

We use the prediction of the ex ante propensity to share to analyze the effect of having the option to view fact-checking separately for Viewers and Nonviewers. The predicted propensity is an estimate of the counterfactual level of sharing in each of the two groups of respondents. We compare the actual sharing intention and the sharing action to the respective ex ante propensities. Consider the actual sharing of alt-facts presented in Panel A.2 of Figure 5. For both Viewers and Nonviewers, the actual sharing rates are significantly lower than the respective ex ante propensities.²¹ The difference between the actual and the predicted sharing intentions (Panel A.1) is also negative for both Viewers and Nonviewers, but for Viewers it is imprecisely estimated.

Panel C of Table 2 presents these results in regression form (with fuller regression output

¹⁹Religious individuals are defined as those who regularly attend religious services, in contrast to those who never attend religious services or who attend only for special ceremonies and festivities.

²⁰The mean outcomes by treatment are slightly different from those presented in Figure 4 since, in Figure 5, the sample is restricted to participants with all pretreatment characteristics nonmissing, which is necessary to calculate the predicted propensity to share, whereas Figure 4 uses the full sample of participants.

²¹Note that the difference between ex ante propensity and the actual outcome in the Imposed Fact-Check treatment group is equal to its average treatment effect.

presented in Online Appendix Table A6). We regress the difference between the actual and the predicted ex ante outcomes measuring the sharing of alt-facts on the dummies for Imposed Fact-Check, Viewers in Voluntary Fact-Check, and Nonviewers in Voluntary Fact-Check, keeping the Alt-Facts treatment as the comparison group. At the bottom of the table, we report the predicted ex ante propensities to share by group, to allow for quantification of the effect of the treatment in percentage terms. Among Viewers, actual sharing decreased compared to the ex ante propensity to share by 39% ($= -0.023/0.0589$) on average as a result of exposure to fact-checking information. Among Nonviewers, actual sharing decreased by 60.2% ($= -0.026/0.0432$) on average as a result of having the option to fact-check, despite the fact that Nonviewers chose not to exercise this option. Thus, as a result of simply being told that fact-checking of the alt-facts is available and can be viewed, the Nonviewers substantially decreased their sharing of alt-facts even though they never saw the fact-checking. (This could be a sign of a rational updating of beliefs about the veracity of news or a purely behavioral effect, such as an increase in salience of accuracy.) The corresponding magnitude of the decrease in sharing intentions compared to its ex ante propensity is 48.5% ($= -0.066/0.136$) for Nonviewers and 13.3% ($= -0.025/0.188$) for Viewers. The effect for Nonviewers is statistically significant, but the effect for Viewers is not. The effects are larger in percentage terms for sharing action than for sharing intention, but, as above, the effects for the intention and for the action are going in the same direction.

Overall, these effects are substantially larger than what could be concluded from a naive comparison of actual sharing across different groups (left-hand-side charts of Panels A.1 and A.2 of Figure 5). Such a comparison would incorrectly assume the same ex ante propensity to share alt-facts between Viewers and Nonviewers.²²

Similarly to the analysis of ex ante propensity to share alt-facts, we calculate the ex ante propensity to share fact-checking (before deciding whether to view) based on all pretreatment characteristics. To do this, we use the actual decisions to share fact-checking in the Imposed Fact-Check treatment group. Columns 3 and 4 of Online Appendix Table A5 present the determinants of sharing of the fact-checking information selected by adaptive LASSO. Some ex ante determinants of sharing fact-checking are similar to those of sharing alt-facts—such as gender and the importance of influence as a motive to share—which suggests that they predict the overall sharing propensity. Other determinants are specific to sharing fact-checking. For instance, having a negative image of the European Union decreases the ex ante propensity of intention to share the fact-checking by 2.6 percentage points and the propensity for sharing action by one percentage point. The propensity to share fact-checking is also associated with self-reported reciprocity: a one-standard-deviation increase in reciprocity increases the propensity of intention to share alt-facts by 1.6 percentage points and the propensity to actually share by 0.8 percentage points.

²²In Online Appendix F, we present a model of the relationship between ex ante propensity to share and the decisions to view the fact-checking and to share ex post; the model derives the conditions under which those who view are ex ante more likely to share.

As with the ex ante propensity to share alt-facts, ex ante propensity to share fact-checking is significantly higher for Viewers than for Nonviewers, as illustrated on the right-hand-side charts of Panels B.1 and B.2 of Figure 5. However, both the actual intention to share fact-checking and the actual sharing of fact-checking among Viewers are even higher. (Nonviewers are, by design, not given an option to share the fact-checking). This raises the question whether some people chose to share both the alt-facts and the fact-checking. It is unclear whether sharing alt-facts and then sharing fact-checking on social media is less damaging than sharing alt-facts only, as alt-facts are more likely to go viral and, more generally, propagate faster and deeper (e.g., Vosoughi et al., 2018). Nevertheless, sharing both alt-facts and fact-checking could limit the propagation of alt-facts, if the followers of the sharer do notice that fact-checking was shared with them together with the alt-facts.

This is why we reproduce the analysis above for three alternative outcomes: share alt-facts only, share fact-checking only, and share both alt-facts and fact-checking. We define these three outcomes both for the sharing intention and for the sharing action. We compute the respective ex ante propensities and compare them to the actual sharing behavior. (Online Appendix Table A7 presents the variable selection by adaptive LASSO.) Figure 6 presents the results. We find that fact-checking dramatically reduces sharing of alt-facts only (alt-facts without fact-checking): the rate of the intention to share “alt-facts only,” i.e., without fact-checking, is 15.3% in the Alt-Facts treatment group, but falls to 7.0% in the Imposed Fact-Check treatment group and to 6.3% among all participants in the Voluntary Fact-Check treatment group. The rates of actual sharing of “alt-facts only” are: 4.8% (Alt-Facts treatment), 2.5% (Imposed Fact-Check treatment), and 2.1% (Voluntary Fact-Check treatment). This implies that our main results (a 27% to 31% reduction in the intention to share alt-facts and a 43% to 45% reduction in sharing alt-facts as a result of the exposure to fact-checking) should be considered as a lower bound for the positive impact of fact-checking. Once we focus on sharing of “alt-facts only” as an outcome, the impact of fact-checking on the sharing intention is 54% to 59% and the impact on the sharing action is 47% to 56%. Further, the mean sharing intention and sharing action of “alt-facts only” are similar for Viewers and Nonviewers, as can be seen on the left-hand-side charts of Panels A.1 and A.2 of Figure 6—despite the fact that Viewers have higher ex ante propensity to share “alt-facts only.” This is because they share *both* alt-facts and fact-checking at least twice as much as the participants in the Imposed Fact-Check treatment group (as illustrated in the left-hand-side charts of Panels C.1 and C.2 of Figure 6). Sharing both alt-facts and fact-checking could be explained either as the choice of individuals who have a strong taste of sharing (regardless of content) or as a choice of those who share alt-facts just to show their friends how outrageous alt-facts are. We cannot separate between the two, as we do not observe the comments people make when they share. (It could also be just a mistake.)

Taken together, the analysis of the interaction between the decision to view fact-checking and the subsequent decision to share alt-facts and fact-checking delivers an optimistic conclusion. Viewing fact-checking reduces sharing of alt-facts and greatly increases sharing of the fact-checking information. Furthermore, the knowledge that fact-checking exists substantially

reduced sharing of alt-facts, even among the participants who choose not to view it. These results explain why the average treatment effects are very similar for the Imposed Fact-Check and Voluntary Fact-Check treatments.

III Interpretation of the Average Treatment Effects

One of our central results is that exposure to fact-checking significantly reduces sharing of alt-facts. In this section, without claiming to be exhaustive, we discuss several possible explanations for this effect. First, it's possible that such exposure changes the internal costs of sharing, and in particular the moral cost of sharing false information. There is an extensive literature on the moral cost of lying (see [Abeler et al., 2019](#), for a survey). If fact-checking increases the perceived likelihood that the news is false, the sharer could face a higher moral cost of sharing something potentially incorrect (see, for instance, [Sperber and Baumard, 2012](#)). Second, it could be that such exposure increases the salience of accuracy and therefore affects the decision to share ([Pennycook et al., 2020b](#)). This interpretation could also explain why Nonviewers in the Voluntary Fact-Check treatment group share alt-facts at a lower rate compared to what is predicted by their characteristics: being offered the possibility to view the fact-checking content makes the issue of accuracy more salient. Third, sharing decisions could be related to social-image concerns (see [Bursztyn and Jensen \(2017\)](#) for a survey and [Bursztyn et al. \(2020b\)](#) for experimental evidence). Sharing false content could negatively affect the sharer's reputation. Finally, fact-checking may reduce sharing of alt-facts if the primary motivation for sharing is influencing the audience: For example, sharing content that has been fact-checked and proven incorrect may result in lower benefits from sharing, because the receivers can get the fact-checking information from other sources and, thus, eventually be less persuaded. These potential interpretations are not mutually exclusive.

Neither our data nor our experimental design allow us to decisively differentiate between these potential mechanisms. However, we can provide some evidence that appears to be at odds with the explanations based on social-image concerns and on intent to influence. In [Table 3](#), we examine whether the exposure to fact-checking has a differential impact on sharing alt-facts on Facebook versus sharing anonymously with 100 participants in another survey experiment, which took place after the one we study in this paper. We create two observations for each participant in our experiment, one for intent to share on Facebook and one for intent to share with other anonymous survey participants. In all columns, we control for socioeconomic characteristics, motivations to share, and behavioral traits. Column 1 shows that individuals share more on Facebook on average. This result may just arise from the fact that individuals were given the possibility to share on Facebook before they were offered the opportunity to share with other participants. We are interested in the differential effects of treatments for sharing on Facebook and with other participants. As reported in Column 2, the coefficients on the interaction terms between being exposed to fact-checking (imposed or voluntary) and

sharing on Facebook are not statistically significant: the difference between sharing on Facebook and sharing with others does not vary across treatments. This result is inconsistent with the explanation related to social-image reputational concerns: Fact-checking equally affects sharing with participants’ Facebook friends (who know the identity of the sharer) and with other experiment participants (who receive anonymous messages). This evidence suggests that, in the context of our experiment, image concerns do not appear to play an important role.²³

In columns 3 to 6 of Table 3, we show that respondents with a larger number of Facebook friends and higher intensity of Facebook use are no more likely to share alt-facts on Facebook than with other participants, controlling for individual characteristics with respondent fixed effects. These findings suggest that the explanations based on the intention to influence the audience are also unlikely: sharing rates are not affected by the size of a potential audience.

A possible concern is whether our results could be driven by Experimenter Demand Effects (EDE), such as the Hawthorne effect (Zizzo, 2010). Throughout the survey, we were careful to be neutral in the way we presented the study to the participants; we believe that it would have been difficult for them to infer what would be the “demand” of the experimenter in our case. Mummolo and Peterson (2018) show that EDEs are not common in studies similar to ours; also, de Quidt et al. (2018) show that EDEs typically are small in magnitude and most often not significantly different from zero, even when experimenters explicitly express their preference to the participants. Nonetheless, in order to test for the importance of EDEs in our experiment, we follow the approach suggested by Dhar et al. (2018). In particular, for each participant, we use pretreatment characteristics to construct the “social desirability score” and check whether the treatment effects are heterogenous across participants with high and low social desirability scores. We define the social desirability score as the first principal component of answers to the following three questions in the survey: “I have worked for a charitable organization,” “I have donated blood,” and “It is important for me not to be considered selfish.” We use these questions to assess how important it is for the participant to appear in a good light to other people, including the experimenters. For each of these questions the participants gave an answer on a scale from 1 to 5 (“never” to “often” for the first two questions, and “it does not fit me” to “it fits me perfectly” for the third question). The three components of the social desirability score are strongly positively correlated and all the factor loadings are positive. We then add the interaction terms between the demeaned social desirability score and treatment dummies to the main specification estimating the average treatment effects (i.e., Panel A of Table 2). The results are reported in the Online Appendix Table A8. We find no evidence of a heterogeneity of treatment effects with respect to the participants’ social desirability score. The coefficients on the interaction terms are never statistically significant. Furthermore, the direct effect of the social desirability score is positive suggesting that people with high scores share Alt-Facts more on average. Overall, we conclude that it is unlikely that there are significant EDEs in our

²³Potentially, consistent behavior between sharing on Facebook and with other participants may be due to an anchoring effect. Yet anchoring does not seem to be an issue in our context, as many participants expressed the intention to share on Facebook but not with other participants, and vice versa: 135 wanted to share only on Facebook, 90 wanted to share only with others, and 167 wanted to share with both groups.

experiment.

IV Discussion of Policy Implications

Circulation of false narratives on social media is an important policy issue. Regardless of the topic—immigration, trade, tax, or health policy—disinformation creates false beliefs that can distort political views and, as a result, voting outcomes (see surveys of empirical evidence in [Guriev and Papaioannou, 2021](#); [Persily and Tucker, 2020](#); [Tucker et al., 2018](#); [Zhuravskaya et al., 2020](#)). Therefore, dissemination of false narratives on social media can have real-world consequences that lead to changes in economic and social policies and even to a rise in hate crime (see, for instance, [Bursztyn et al., 2020a](#); [Mueller and Schwarz, 2020, 2021](#)).

An unregulated equilibrium in which social media users consume—and act based upon—false information is inefficient. Like any costly information, fact-checking is a public good. The individual return on a fact-checker’s costly effort is much lower than the social return. Without a policy intervention, investment in fact-checking is, therefore, inefficiently low in equilibrium.²⁴

This market failure can be addressed by fiscal or regulatory interventions: governments could either subsidize fact-checking directly or require social media platforms to support it. However, the design of an effective policy response to online dissemination of disinformation cannot be limited to basic economic analysis of individual incentives to produce fact-checking. Policymakers also have to account for the psychology of consuming alt-facts and fact-checking. The business model of social media is based on keeping users’ attention and creating addiction to the platform ([Singer and Brooking, 2018](#), ch. 8). This results in the emergence of the “dark psychology of social networks” [Haidt and Rose-Stockwell \(2019\)](#); [Tufekci \(2018\)](#): as alt-facts attract more attention, they are more likely to keep users engaged. Therefore, social media platforms are interested in their wide circulation (see [Allcott and Gentzkow \(2017\)](#); [Silverman \(2016\)](#); [Vosoughi et al. \(2018\)](#) for evidence on circulation of political false narratives on Facebook and Twitter). These psychological effects further increase the social inefficiency of an unregulated equilibrium.

Our results suggest policies that can limit the circulation of alternative facts. First, we show that exposing social-media users to fact-checking information substantially decreases sharing of alt-facts. Second, we show that this reduction is as large when the fact-checking is imposed on users as it is when users are given the option to view it. This is important, because “voluntary” exposure to fact-checking is much easier to implement than “imposed” fact-checking, for practical, legal, and ethical reasons. Third, we find that each additional click required to confirm sharing reduces the number of sharers substantially.

Our findings imply that policies offering users an opportunity to fact-check suspicious statements and requiring additional clicks to confirm the sharing of unchecked or false posts could be highly effective. These policies are also realistic: platforms have recently started to

²⁴See [Grossman and Stiglitz \(1980\)](#) for a theory of inefficiency of equilibrium with decentralized acquisition of information.

implement them, albeit on a limited scale and as a result of a public pressure rather than government regulation.

As we noted in the introduction, since May 2020, Twitter has been adding a blue link to fact-checking information (essentially equivalent to our Voluntary Fact-Checking treatment) to each of Donald Trump’s misleading and false tweets. Twitter understands the importance of each additional click: since June 2020, Twitter has been testing a feature requiring users to click on a link before being able to share it (Montti, 2020). In May 2020, Twitter also introduced an additional click for users wishing to see Trump’s “when looting starts, shooting starts” tweet.²⁵

Facebook also has implemented several measures to counter propagation of false news. Since April 2017, Facebook has partnered with fact-checking organizations around the world. This growing alliance now includes more than 50 partners, such as the International Fact-Checking Network, PolitiFact.com, Factly, Agence France Presse, *Le Monde*, and *Libération*. Facebook’s algorithm identifies suspicious posts and submits them to the partners. Partners choose which posts to check, if any, and are paid by Facebook for providing fact-checking. If a post is identified as false, its content is made less visible and Facebook adds a link to the fact-checking article or to other relevant resources.

There is a growing concern that the social media platforms’ self-imposed fact-checking programs are insufficient in scale and scope. As Singer and Brooking (2018, p. 247) argue, “... never before has so much, posted by so many, has been moderated by so few.” One of Facebook’s fact-checking partners, Snopes.com, left, whereas several others have refused to join the program, pointing to low financial compensation (Madrigal, 2019). Bengani and Karbal (2020) estimate that in 2019 Facebook spent about \$1.2 million on fact-checking, or about 0.001% of its \$71 billion revenue (and about 10^{-6} of its market capitalization). The scope of its fact-checking has also been limited: Facebook has decided not to fact-check politicians’ speeches. In addition, Facebook often chooses not to follow fact-checkers’ advice. Bengani and Karbal (2020) find that the company has failed to consistently label content flagged by its fact-checking partners. During a five-day period in October 2020, Facebook’s ten U.S. fact-checking partners debunked over 70 claims. The authors identify over 1,100 posts on Facebook and Instagram containing these debunked falsehoods; among them, less than one half were flagged by Facebook as false.

It seems clear that self-regulation is unlikely to restore the social optimum, because the platforms’ incentives—even after accounting for reputation concerns—are not aligned with the social optimum. First, platforms are for-profit businesses that are not interested in spending resources on social goals beyond their shareholder value. Second, fact-checking and reduced circulation of alt-facts also leads to lower user engagement, which undermines advertising revenue.

²⁵The fact that, in January 2021, Twitter first banned resharing of Donald Trump’s tweets, then permanently closed his account is also consistent with our results. We show that fact-checking and an additional click substantially reduce sharing of alt-facts but do not eliminate such sharing completely. The social cost of propagation of some alt-facts originating from important opinion makers may be so large that blunter instruments, such as a sharing ban or an outright account closure, may be optimal.

Our work suggests that an optimal policy intervention should include public support for independent fact-checking organizations and a requirement that platforms partner with them. Regulators should also require the social media platforms to label as false all disinformation identified by fact-checkers and provide links to the relevant fact-checking. Finally, the platforms should introduce additional clicks for users intending to share content that was not fact-checked or that was fact-checked and proven false.

V Conclusions

Our paper makes several contributions to the literature. First, using real-world decisions to share content on social networks, we show that exposing social media users to fact-checking information significantly (by 43% to 45%) decreases their sharing of alternative facts. Second, this reduction is the same whether users are forced to read fact-checking information or are simply given the opportunity to do so. Third, we show that users who choose to view the fact-checking are ex ante more likely to share both alt-facts and fact-checking information. These results suggest an important role that fact-checking can play in limiting propagation of false news on social media. Fourth, we show that additional clicks required to share content on social media substantially reduce sharing.

As propagation of alt-facts on social media imposes a negative externality, there is a rationale for policy interventions that reduce sharing of alt-facts. Our analysis helps identify potential solutions. First, we show that providing users the option to view fact-checking information substantially reduces sharing of false statements. Offering voluntary fact-checking is much easier to implement than imposing fact-checking on users. Second, our results imply that sharing is highly sensitive to even very small nonmonetary costs, such as introducing additional clicks. Therefore, there is a scope for regulation that demands additional clicks, for example, asking users to confirm that they are aware that a particular statement is likely to be false. Optimal design of such interventions is an important avenue for future research and public debate.

References

- Abeler, Johannes, Daniele Nosenzo, and Collin Raymond**, “Preferences for Truth-Telling,” *Econometrica*, 2019, *87*, 1115–1153.
- Abreu, Luis and Doh-Shin Jeon**, “Homophily in Social Media and News Polarization,” *working paper*, 2019.
- Allcott, Hunt and Matthew Gentzkow**, “Social Media and Fake News in the 2016 Election,” *Journal of Economic Perspectives*, 2017, *31* (2), 211–36.
- Altay, Sacha, Emma De Araujo, and Hugo Mercier**, “If this account is true, it is most enormously wonderful: Interestingness-if-true and the sharing of true and false news,” 2020.
- Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya**, “Facts, alternative facts, and fact checking in times of post-truth politics,” *Journal of Public Economics*, 2020, *182*, 104–123.
- Bénabou, Roland, Armin Falk, and Jean Tirole**, “Narratives, Imperatives, and Moral Persuasion,” *NBER working paper*, 2019.
- Bengani, Priyanjana and Ian Karbal**, “Five Days of Facebook Fact-Checking,” *Columbia Journalism review*, 2020.
- Bloch, Francis, Gabrielle Demange, and Rachel Kranton**, “Rumors and social networks,” *International Economic Review*, 2018, *59* (2), 421–448.
- Bursztyn, Leonardo and Robert Jensen**, “Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure,” *Annual Review of Economics*, 2017, *9*, 131–153.
- , **Georgy Egorov, Ruben Enikolopov, and Maria Petrova**, “Social Media and Xenophobia: Evidence from Russia,” *CEPR Discussion Paper 14877*, 2020.
- , **Ingar Haaland, Aakaash Rao, and Christopher Roth**, “Disguising Prejudice: Popular Rationales as Excuses for Intolerant Expression,” *Mimeo, University of Chicago*, 2020.
- Campbell, A., M. Leister, and Y. Zenou**, “Social media and polarization,” *working paper*, 2019.
- Conger, Kate**, “A Look Inside Twitter’s Move To Flag Trump,” *New York Times*, 2020, *May 30*, A1. <https://www.nytimes.com/2020/05/30/technology/twitter-trump-dorsey.html> (accessed on Jan 14, 2021).
- , “Twitter Has Labeled 38% of Trump’s Tweets Since Tuesday,” *New York Times*, 2020, *November 5*. <https://www.nytimes.com/2020/11/05/technology/donald-trump-twitter.html> (accessed on Jan 14, 2021).
- **and Davey Alba**, “Trump Posts On Twitter Are Labeled For Falseness,” *New York Times*, 2020, *May 27*, B1. <https://www.nytimes.com/2020/05/26/technology/twitter-trump-mail-in-ballots.html> (accessed on Jan 14, 2021).
- de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth**, “Measuring and Bounding Experimenter Demand,” *American Economic Review*, 2018, *108*(11), 3266–3302.

- Dhar, Diva, Tarun Jain, and Seema Jayachandran**, “Reshaping Adolescents’ Gender Attitudes: Evidence from a School-Based Experiment in India,” Working Paper 25331, National Bureau of Economic Research December 2018.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde**, “Representative Trust and Reciprocity: Prevalence and determinants,” *Economic Inquiry*, 2008, 46, 84–90.
- Fazio, Lisa**, “Pausing to consider why a headline is true or false can help reduce the sharing of false news,” *The Harvard Kennedy School Misinformation Review*, 2020.
- Grossman, Sanford J. and Joseph E. Stiglitz**, “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 1980, 70(3), 393–408.
- Guess, Andrew, Jonathan Nagler, and Joshua Tucker**, “Less than you think: Prevalence and predictors of fake news dissemination on Facebook,” *Science Advances*, 2019, 5.
- Guriev, Sergei and Elias Papaioannou**, “The Political Economy of Populism,” *Journal of Economic Literature*, 2021, *forthcoming*.
- , **Emeric Henry, and Ekaterina Zhuravskaya**, “Sharing fake news,” May 2019. AEA RCT Registry. <https://doi.org/10.1257/rct.4060-1.0>.
- Haidt, Jonathan and Tobias Rose-Stockwell**, “The Dark Psychology of Social Networks: Why it feels like everything is going haywire,” *The Atlantic*, 2019. <https://www.theatlantic.com/magazine/archive/2019/12/social-media-democracy/600763/> accessed on July 22, 2020.
- Henry, Emeric and Jan Sonntag**, “Measuring image concern,” *Journal of Economic Behavior and Organization*, 2019, 160, 19–39.
- , **Ekaterina Zhuravskaya, and Sergei Guriev**, “Data and Code for: Checking and Sharing Alt-Facts,” 2021. *American Economic Association* [publisher], Inter-university Consortium for Political and Social Research [distributor]. <http://doi.org/10.3886/E140161V1>.
- Madrigal, Alexis**, “This Is How Much Fact-Checking Is Worth to Facebook: More than nothing, but not much more,” *The Atlantic*, 2019.
- Montti, Roger**, “Twitter Will Check if Articles Are Read Before Sharing,” *Search Engine Journal*, 2020, June 11. <https://www.searchenginejournal.com/twitter-sharing-prompt/371825/> accessed on January 14, 2021.
- Mosleh, Mohsen, Gordon Pennycook, and David Rand**, “Self-reported willingness to share political news articles in online surveys correlates with actual sharing on Twitter,” *PLoS ONE*, 2020, 15(2): e0228882.
- Mueller, Karsten and Carlo Schwarz**, “From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment,” Technical Report, Mimeo, Warwick University 2020.
- and – , “Fanning the Flames of Hate: Social Media and Hate Crime,” *Journal of the European Economic Association*, 2021, *Forthcoming*.
- Mummolo, Jonathan and Erik Peterson**, “Demand effects in survey experiments: An empirical assessment,” *American Political Science Review*, 2018, pp. 1–13.

- Nyhan, Brendan, Ethan Porter, Jason Reifler, and Thomas Wood**, “Taking Fact-Checks Literally But Not Seriously? The Effects of Journalistic Fact-Checking on Factual Beliefs and Candidate Favorability,” *Political Behavior*, 2019.
- Pennycook, Gordon, Adam Bear, Evan Collins, and David Rand**, “The Implied Truth Effect: Attaching Warnings to a Subset of FakeNews Headlines Increases Perceived Accuracy of HeadlinesWithout Warnings,” *Management Science*, 2020.
- , **Jonathon Mcphetres, Yunhao Zhang, and David Rand**, “Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention.,” *Psychological Science*, 2020, *31*, 770–780.
- , **Ziv Epstein, Mohsen Mosleh, Antonio Arechar, Dean Eckles, and David Rand**, “Understanding and reducing the spread of misinformation online,” 2019.
- Persily, Nathaniel and Joshua A. Tucker**, *Social Media and Democracy*, Cambridge University Press, 2020.
- Rushton, J. P., R. D. Chrisjohn, and G. C. Fekken**, “The altruistic personality and the self-report altruism scale.,” *Personality and Individual Differences*, 1981, *2*, 293–302.
- Silverman, Craig**, “This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook,” *BuzzFeed*, 2016, *November 16*. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook> (accessed on Jan 14, 2021).
- Singer, P.W. and Emerson T. Brooking**, *LikeWar: The Weaponization of Social Media*, Mariner Books, 2018.
- Sperber, Dan and Nicolas Baumard**, “Moral Reputation: An Evolutionary and Cognitive Perspective,” *Mind and Language*, 2012, *27*, 495–518.
- Swire, Briony, Adam Berinsky, Lewandowsky, and Ullrich Ecker**, “Processing political misinformation: comprehending the Trump phenomenon,” *Royal Society Open Science*, 2017, *4* (3).
- Töller, Annette Elizabeth**, “Measuring and Comparing the Europeanization of National Legislation: A Research Note,” *Journal of Common Market Studies*, 2010, *48*.
- Tucker, Joshua A., Andrew Guess, Pablo Barbera, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan**, “Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature,” Technical Report, Hewlett Foundation 2018.
- Tufekci, Zeynep**, “How social media took us from Tahrir Square to Donald Trump,” Technical Report, MIT Technology Review 2018.
- Vosoughi, Soroush, Deb Roy, and Sinan Aral**, “The spread of true and false information online,” *Science*, 2018, *359*, 1146–1151.
- Yaqub, Waheeb, Otari Kakhidze, Morgan Brockman, Nasir Memon, and Sameer Patil**, “Effects of Credibility Indicators on Social Media News Sharing Intent,” *CHI ’20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov, “Political Effects of the Internet and Social Media,” *Annual Review of Economics*, 2020, *12*, 415–438.

Zizzo, Daniel John, “Experimenter Demand Effects in Economic Experiments,” *Experimental Economics*, mar 2010, *13* (1), 75–98.

Zou, Hui, “The Adaptive Lasso and Its Oracle Properties,” *Journal of the American Statistical Association*, dec 2006, *101* (476), 1418–1429.

Figure 1: Questions about the Sharing Intention and the Actual Sharing

Screen 1: measurement of “Want to share alt-facts” variable

Do you want to share this article “RN: the European Union controls 87% of our laws and wants to open the door to 50 million of immigrants” containing the declarations made by Marine Le Pen and Jordan Bardella on your Facebook page?

- Yes
- No
- Do not wish to answer

Screen 2: measurement of “Share alt-facts” variable

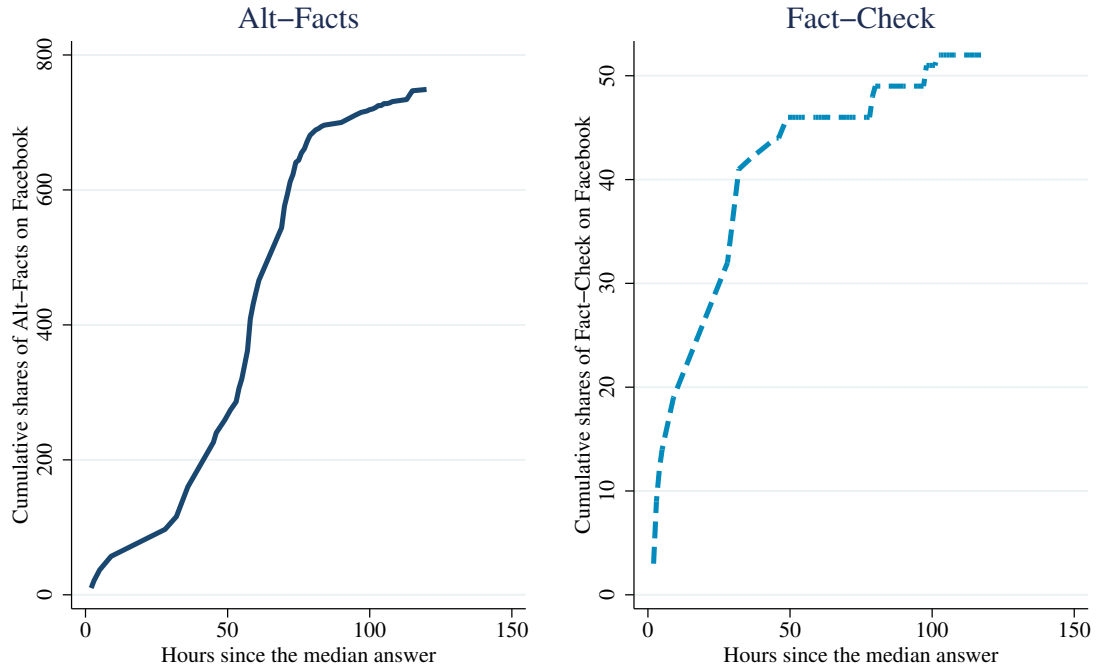
You can share the article you just read on your Facebook page by clicking on the share button (if you are not already logged-on and you want to share the content, you can either log-on now on a separate page, or you will need to log-on after you click on the share button)



Note: The figure reports the english translation of the questions about the sharing intention and the actual sharing in our experiment. The original formulation in French is presented in Online Appendix B.

Figure 2: Aggregate Sharing on Facebook

Cumulative propagation of Alt-Facts and Fact-Check on Facebook
(including resharing by users outside the experiment)



Note: The cumulative shares of alt-facts (in the right panel) and of fact-checking (in the left panel) counts the cumulative number of shares both by the participants in our survey-experiment as well as other Facebook users. We define date 0 as the date at which the median respondent ends the survey.

Figure 3: The Effect of the Number of Clicks Needed to Share

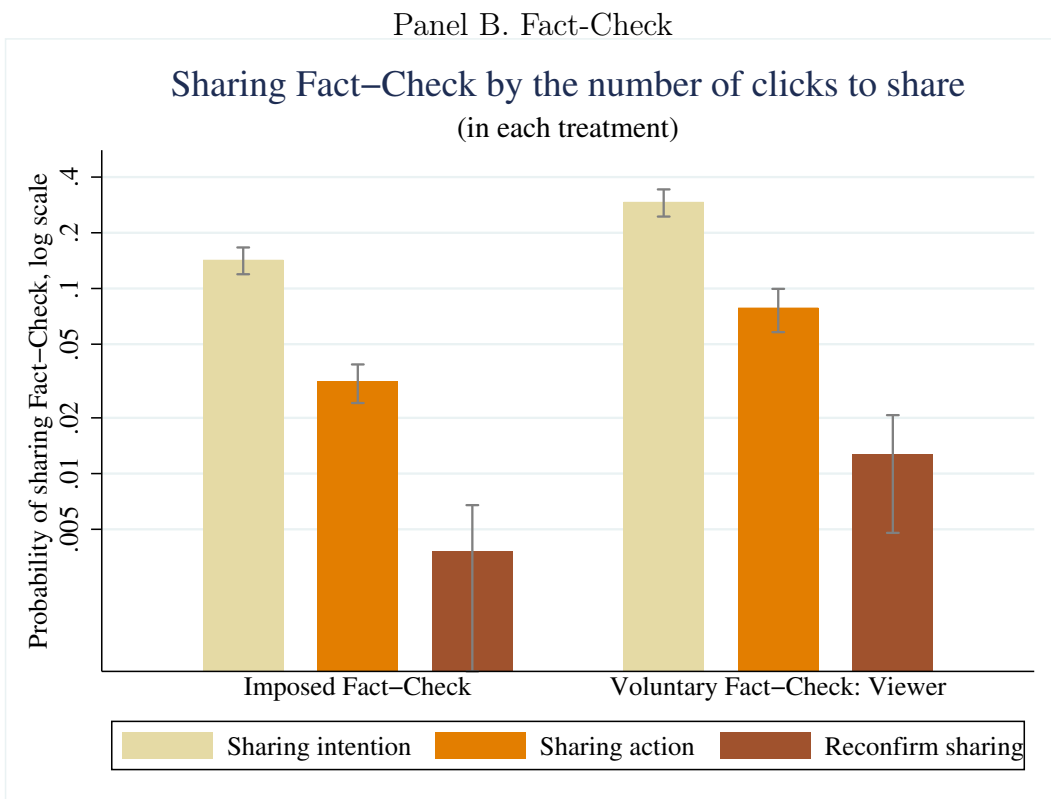
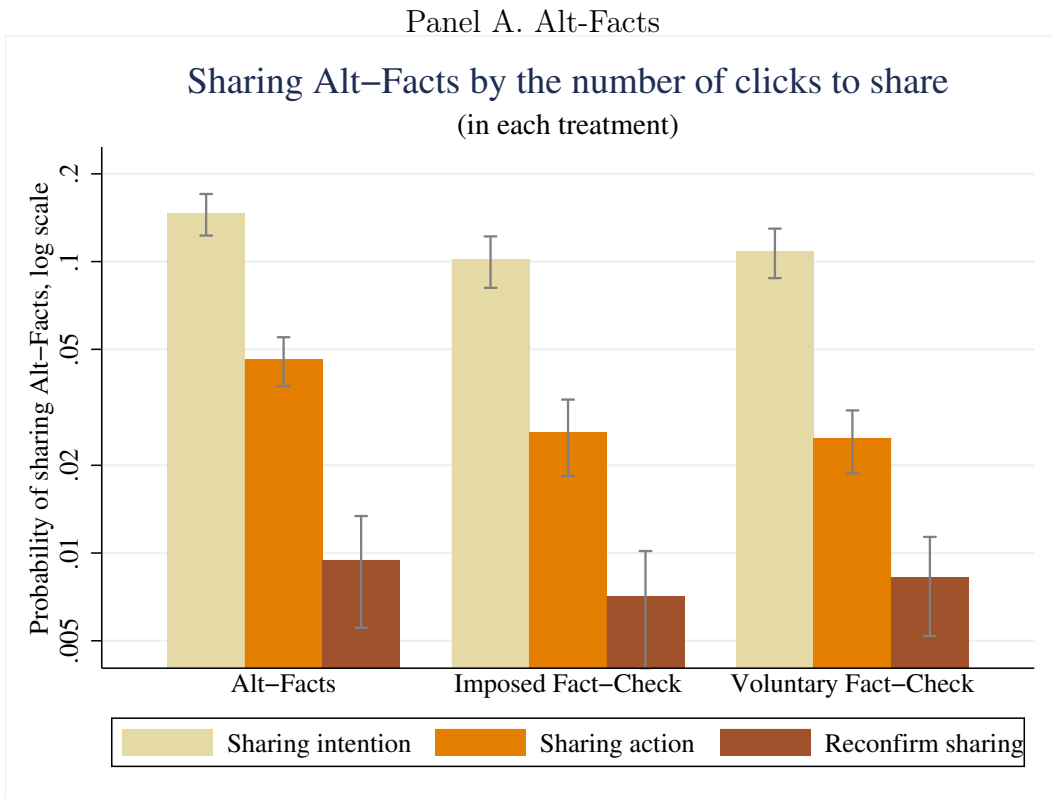
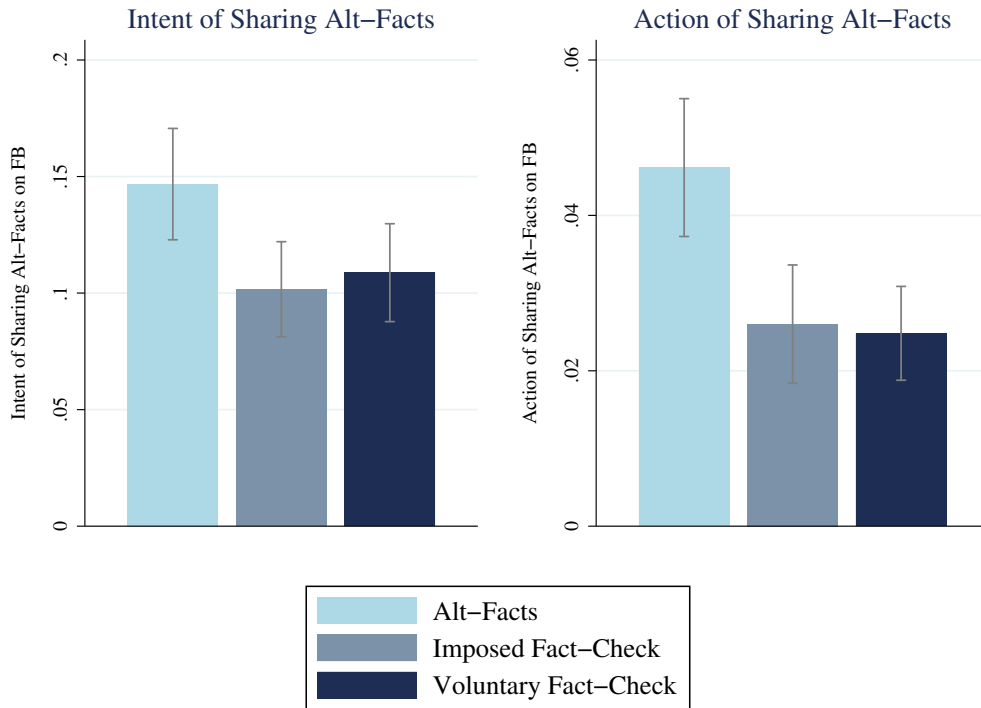


Figure 4: Average Treatment Effects

Panel A. Share Alt-Facts on Facebook



Panel B. Share Fact-Check on Facebook

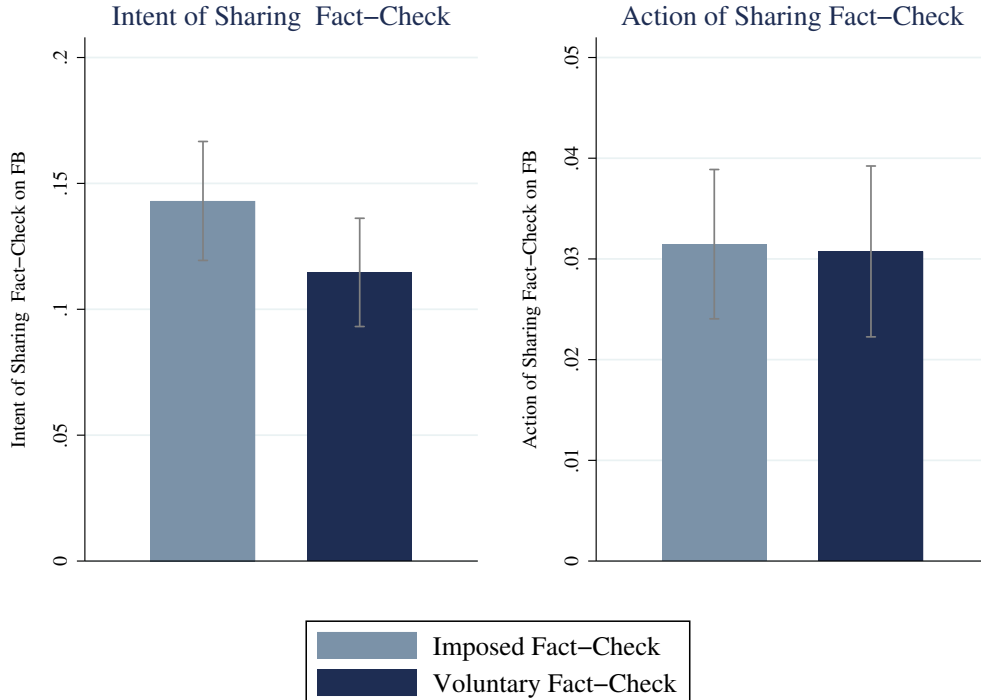
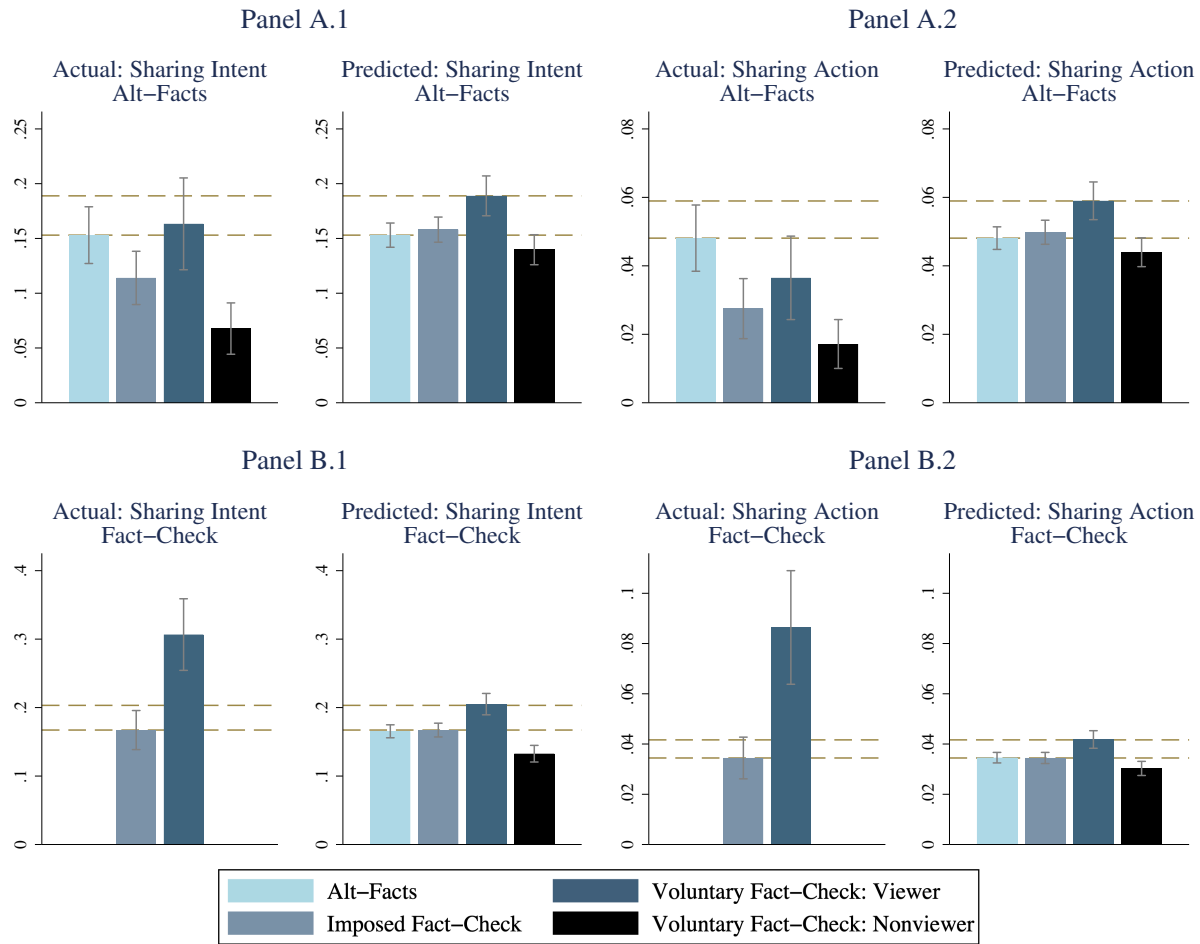
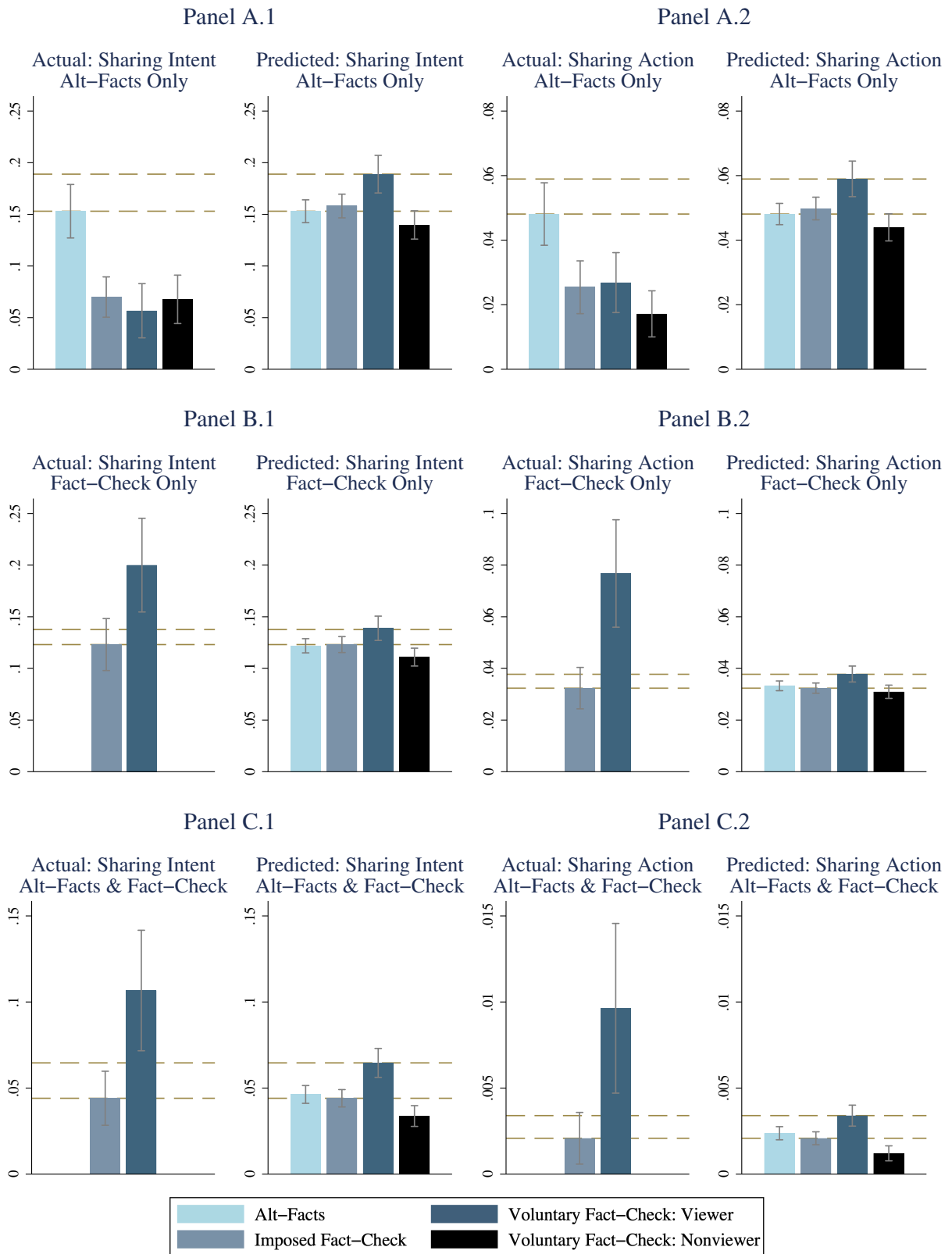


Figure 5: Actual vs. Predicted Sharing of Alt-Facts and Fact-Check, by Treatment



Note: The top horizontal line indicates the predicted level of sharing for viewers in the voluntary fact-check treatment; the bottom horizontal line indicates the predicted level of sharing in alt-facts treatment in Panels A.1 and A.2 (the top row) and in imposed fact-check treatment in the Panels B.1 and B.2 (bottom row).

Figure 6: Actual vs. Predicted Sharing of Alt-Facts Only, Fact-Check Only, and Both Alt-Facts and Fact-Check, by Treatment



Note: The top horizontal line indicates the predicted level of sharing for viewers in the voluntary fact-check treatment; the bottom horizontal line indicates the predicted level of sharing in alt-facts treatment in Panels A.1 and A.2 (top row) and in imposed fact-check treatment in Panels B.1, B.2, C.1, and C.2 (the middle and bottom row).

Table 1: Balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Treatments:						Difference in means:	
	(a) Alt-Facts		(b) Imposed Fact-Check		(c) Voluntary Fact-Check		(a) vs (b)	(a) vs (c)
	mean	sd	mean	sd	mean	sd	p	p
Male	0.50	0.50	0.50	0.50	0.50	0.50	(0.94)	(0.94)
Education level (1-9)	5.34	2.52	5.47	2.46	5.10	2.57	(0.29)	(0.05)
Born in France	0.95	0.22	0.96	0.21	0.96	0.19	(0.61)	(0.24)
Father born in France	0.84	0.37	0.87	0.34	0.87	0.33	(0.08)	(0.05)
Mother born in France	0.86	0.35	0.90	0.31	0.89	0.31	(0.03)	(0.04)
Age	43.51	14.81	43.57	14.89	45.98	15.41	(0.94)	(0.00)
Live in village	0.28	0.45	0.25	0.44	0.26	0.44	(0.25)	(0.40)
Live in city	0.54	0.50	0.54	0.50	0.52	0.50	(0.86)	(0.47)
Married	0.36	0.48	0.39	0.49	0.43	0.50	(0.20)	(0.00)
Single	0.26	0.44	0.24	0.42	0.23	0.42	(0.26)	(0.17)
Income (1-10)	4.41	2.54	4.51	2.53	4.59	2.59	(0.39)	(0.15)
Children	0.66	0.47	0.66	0.47	0.69	0.46	(0.93)	(0.19)
Frequency of FB use	3.57	0.74	3.52	0.79	3.52	0.79	(0.16)	(0.16)
Log (FB friends+1)	4.38	1.44	4.37	1.45	4.23	1.56	(0.90)	(0.04)
Often share on FB	0.50	0.50	0.46	0.50	0.48	0.50	(0.14)	(0.37)
Reason to share: interest	0.71	0.21	0.70	0.21	0.71	0.22	(0.29)	(0.93)
Reason to share: influence	0.47	0.24	0.46	0.24	0.49	0.25	(0.48)	(0.30)
Reason to share: image	0.55	0.28	0.51	0.26	0.54	0.27	(0.01)	(0.52)
Reason to share: reciprocity	0.63	0.25	0.60	0.25	0.61	0.25	(0.02)	(0.17)
Source news: TV	0.41	0.49	0.45	0.50	0.45	0.50	(0.08)	(0.09)
Source news: Internet	0.35	0.48	0.32	0.47	0.30	0.46	(0.26)	(0.03)
Self-reported altruism	0.45	0.20	0.44	0.18	0.44	0.20	(0.13)	(0.13)
Self-reported reciprocity	0.79	0.18	0.78	0.18	0.78	0.19	(0.71)	(0.42)
Self-reported image importance	0.74	0.24	0.72	0.25	0.73	0.25	(0.11)	(0.13)
Gave money to charity	0.49	0.27	0.48	0.26	0.48	0.27	(0.44)	(0.35)
Gave money to homeless	0.41	0.25	0.40	0.24	0.39	0.26	(0.30)	(0.09)
Worked for charity	0.44	0.29	0.41	0.27	0.41	0.27	(0.06)	(0.06)
Donated blood	0.48	0.31	0.47	0.31	0.48	0.32	(0.53)	(0.98)
Religious	0.20	0.40	0.21	0.41	0.22	0.41	(0.44)	(0.31)
Religion: catholic	0.39	0.49	0.45	0.50	0.46	0.50	(0.02)	(0.01)
Religion: muslim	0.05	0.22	0.05	0.21	0.04	0.21	(0.82)	(0.64)
Feel closer to European gov.	0.12	0.32	0.11	0.32	0.11	0.31	(0.81)	(0.70)
Feel closer to French gov.	0.50	0.50	0.48	0.50	0.52	0.50	(0.51)	(0.27)
Negative Image EU (1-5)	3.24	1.11	3.24	1.05	3.21	1.10	(0.98)	(0.53)
Voted Le Pen, 1st round 2017	0.24	0.43	0.22	0.41	0.22	0.42	(0.20)	(0.24)
Voted Macron, 1st round 2017	0.14	0.35	0.14	0.35	0.19	0.39	(0.94)	(0.01)
Voted Le Pen, 2nd round 2017	0.27	0.45	0.26	0.44	0.26	0.44	(0.50)	(0.65)
Observations	845		846		846		1691	1691

Note: First three columns present mean values and standard deviations by treatment groups for the main pretreatment variables. The last two columns present the p-values for the test of the equality of means between the Alt-Facts treatment group and the two Fact-Check treatment groups. We present the omnibus test of randomization quality for all pre-treatment characteristics in Online Appendix Table A1.

Table 2: Average Treatment Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Average treatment effect on sharing alt-facts								
Dependent variable:	Intent to Share Alt-Facts on Facebook				Action of Sharing Alt-Facts on Facebook			
Imposed Fact-Check	-0.045 (0.016)	-0.045 (0.016)	-0.049 (0.017)	-0.050 (0.017)	-0.020 (0.006)	-0.020 (0.006)	-0.021 (0.006)	-0.021 (0.007)
Voluntary Fact-Check	-0.038 (0.016)	-0.038 (0.016)	-0.040 (0.017)	-0.052 (0.017)	-0.021 (0.005)	-0.021 (0.005)	-0.024 (0.006)	-0.026 (0.006)
Observations	2,537	2,537	2,265	2,078	2,537	2,537	2,265	2,078
R-squared	0.004	0.010	0.110	0.166	0.008	0.010	0.073	0.101
Mean DV, Alt-Facts Treatment	0.147	0.147	0.155	0.161	0.0462	0.0462	0.0492	0.0509
p-value, Imposed=Voluntary	0.635	0.634	0.530	0.899	0.812	0.812	0.579	0.379
Panel B: Average treatment effect on sharing fact-checking								
Dependent variable:	Intent to Share Fact-Check on Facebook				Action of Sharing Fact-Check on Facebook			
Voluntary Fact-Check	-0.028 (0.016)	-0.028 (0.016)	-0.041 (0.017)	-0.039 (0.018)	-0.001 (0.006)	-0.001 (0.006)	-0.001 (0.006)	0.004 (0.007)
Observations	1,692	1,692	1,517	1,388	1,692	1,692	1,517	1,388
R-squared	0.002	0.038	0.100	0.131	0.000	0.004	0.038	0.073
Mean DV, Imposed Fact-Check	0.143	0.143	0.158	0.160	0.0315	0.0315	0.0337	0.0320
Panel C: Actual compared to predicted sharing for viewers and nonviewers								
Dependent variable:	Difference b/w actual and predicted: Intent to Share Alt-Facts on Facebook				Difference b/w actual and predicted: Action of Sharing Alt-Facts on Facebook			
Imposed Fact-Check	-0.046 (0.016)	-0.045 (0.016)	-0.052 (0.017)	-0.051 (0.018)	-0.021 (0.006)	-0.021 (0.006)	-0.022 (0.007)	-0.021 (0.007)
Voluntary Fact-Check: Viewer	-0.025 (0.023)	-0.028 (0.023)	-0.028 (0.024)	-0.025 (0.024)	-0.023 (0.008)	-0.022 (0.008)	-0.022 (0.008)	-0.020 (0.008)
Voluntary Fact-Check: Nonviewer	-0.066 (0.017)	-0.063 (0.017)	-0.074 (0.018)	-0.072 (0.019)	-0.026 (0.006)	-0.027 (0.006)	-0.032 (0.007)	-0.030 (0.007)
Observations	2,219	2,219	2,051	2,018	2,349	2,349	2,112	2,078
R-squared	0.007	0.011	0.038	0.069	0.010	0.016	0.046	0.063
Mean Dep. var., Alt-Facts T	-0.0014	-0.0014	0.0036	0.0023	0.0006	0.0006	0.0027	0.0019
Mean predicted Alt-Facts T	0.150	0.150	0.156	0.159	0.0474	0.0474	0.0483	0.0490
Mean Predict. Imposed Fact-Check T	0.158	0.158	0.159	0.159	0.0467	0.0467	0.0479	0.0480
Mean Predict., Voluntary T, Viewer	0.188	0.188	0.190	0.190	0.0589	0.0589	0.0592	0.0592
Mean Predict., Voluntary T, Nonviewer	0.136	0.136	0.144	0.144	0.0432	0.0432	0.0442	0.0442
All panels:								
Strata controls		✓	✓	✓		✓	✓	✓
Demographic controls & facebook use			✓	✓			✓	✓
All pretreatment characteristics				✓				✓

Note: The comparison group in Panel A is Alt-Facts treatment. The comparison group in Panel B is Imposed Fact-Check treatment. Strata controls include dummies for gender and middle-level and low-level education. The set of demographic & facebook use controls is as follows: age, age squares, income (1-10), family status (dummies for married and single), number of children, size of the place of living (dummies for village and town), dummy for attending religious services, religion (dummies for catholic, muslim, and no religion), dummies for each of the 9 levels of education, frequency of Facebook use, dummy for sharing often on Facebook, the log of 1+ the of Facebook friends, a dummy for having voted for Marin Le Pen in the second round of 2017 elections, negative image EU. The list of all pretreatment characteristics is given in Online Appendix Table A1. Most importantly, it includes self-reported altruism, reciprocity, and image importance, as well as self-reported reasons to share on Facebook (interest, influence, image, reciprocity). Robust standard errors are in parentheses.

Table 3: Sharing on Facebook vs. sharing with other participants

	(1)	(2)	(3)	(4)	(5)	(6)
	Intent to share Alt-Facts either on Facebook or with other participants					
Intent to Share on Facebook	0.018 (0.006)	0.020 (0.017)	0.018 (0.006)	0.015 (0.008)	0.011 (0.018)	0.014 (0.008)
Imposed Fact-Check	-0.050 (0.013)	-0.056 (0.014)				
Voluntary Fact-Check	-0.029 (0.014)	-0.020 (0.016)				
Intent to Share on Facebook × × Imposed Fact-Check		0.011 (0.022)				
Intent to Share on Facebook × × Voluntary Fact-Check		-0.018 (0.023)				
Intent to Share on Facebook × × Number of Facebook friends >100				0.007 (0.012)		
Intent to Share on Facebook × × log(Facebook friends+1)					0.001 (0.004)	
Intent to Share on Facebook × × Share on Facebook often						0.009 (0.012)
Observations	5,074	5,074	5,074	5,074	4,970	5,074
R-squared	0.005	0.006	0.004	0.004	0.003	0.004
Respondent FEs			✓	✓	✓	✓

Note: We use all survey participants and create two observation per participant, one for the decision to share on Facebook and one for the decision to share with others. The dependent variable takes value 1 if the alt-facts were shared. The variable “Share on Facebook” takes the value 1 if the observation corresponds to the sharing decision on Facebook. Column (3) to (6) include individual FEs. Robust standard errors are in parentheses.

Online Appendix

for

“Checking and Sharing Alt-Facts”

by Emeric Henry

©

Ekaterina Zhuravskaya

©

Sergei Guriev

Content:

A. Additional empirical results

B. Questions on sharing intention and actual sharing in our experiment and in related studies

C. Text of the treatments

D. Complete questionnaire

E. Definition of variables

F. Theoretical model

A Additional empirical results

Figure A1: The distributions of values in the variables “Action of sharing alt-facts on Facebook” and “Action of sharing fact-check on Facebook” across respondents

Sharing on Facebook among respondents, who expressed the intent to share

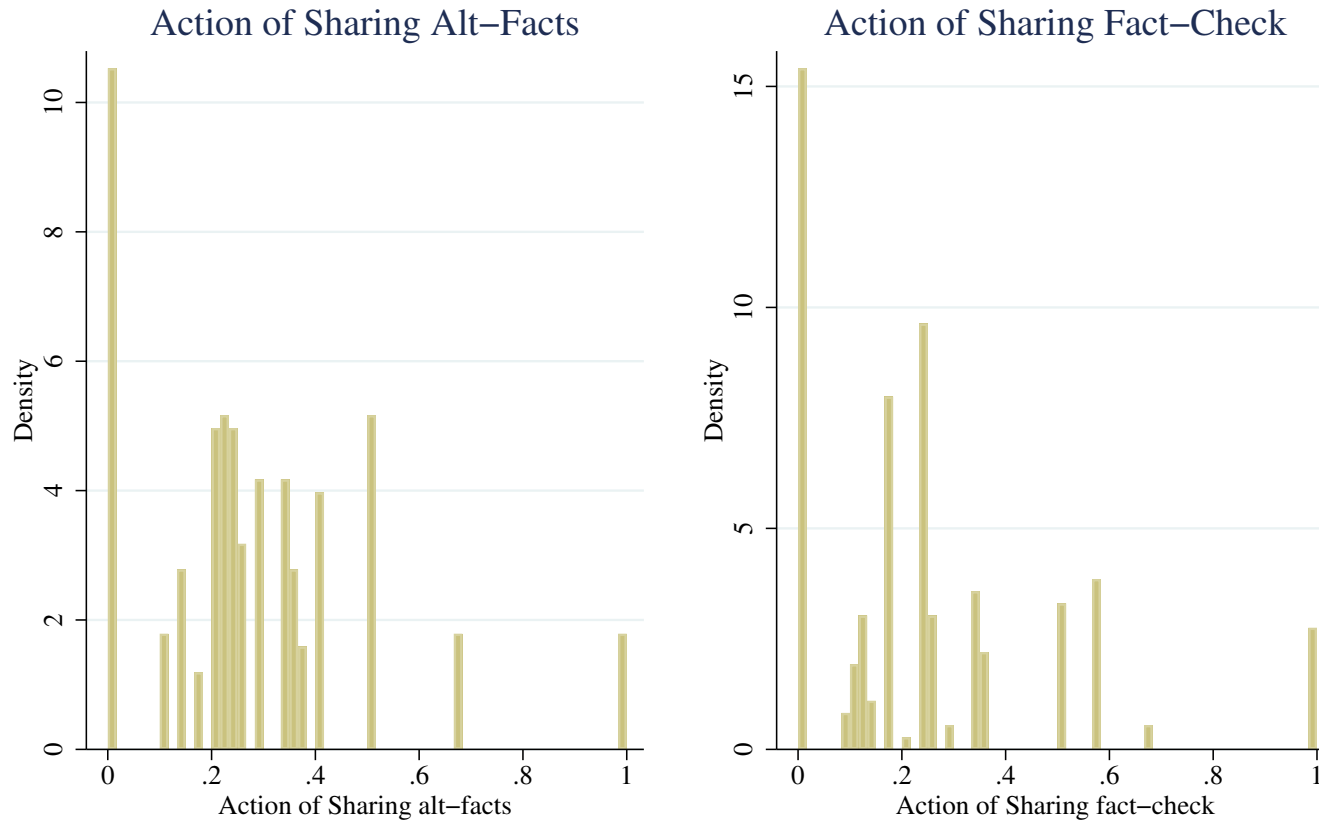


Figure A2: Viewing fact-checking in the Voluntary Fact-Check treatment and the ex ante predictors of sharing alt-facts

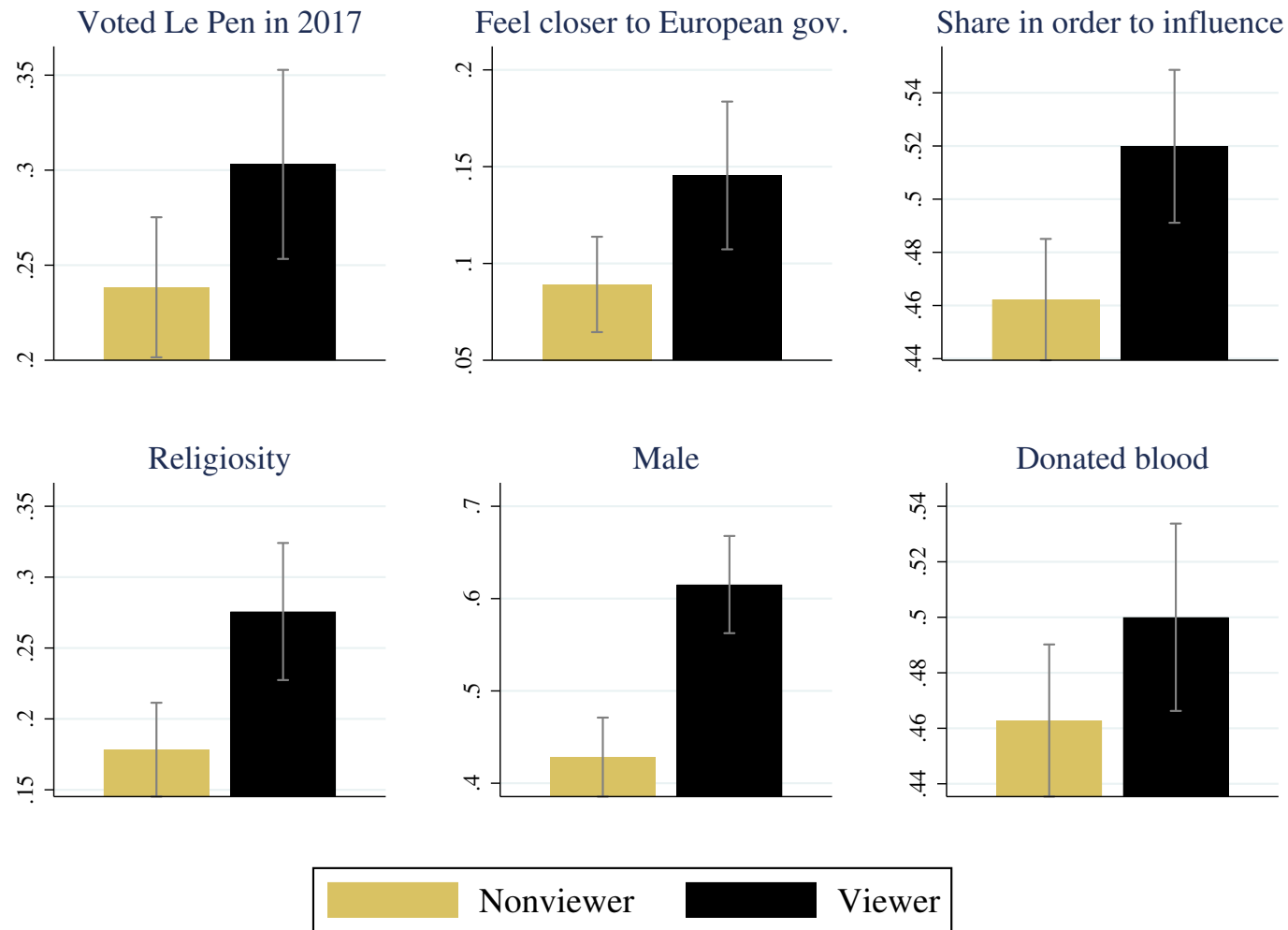


Table A1: Omnibus test of randomization quality, multinomial logistic regression

	(1)	(2)		(1) continued	(2) continued		(1) continued	(2) continued
Outcome (treatment):	Imposed FC (vs. Alt-Facts)	Voluntary FC		Imposed FC (vs. Alt-Facts)	Voluntary FC		Imposed FC (vs. Alt-Facts)	Voluntary FC
Mid education	-0.155 (0.184)	-0.083 (0.180)	Frequency of FB use	-0.000 (0.085)	0.046 (0.089)	Gave money to homeless	-0.159 (0.341)	-0.059 (0.340)
High education	-0.059 (0.184)	-0.026 (0.182)	Log (FB friends+1)	0.017 (0.046)	-0.019 (0.047)	Worked for charity	-0.454 (0.329)	-0.170 (0.314)
Old	0.098 (0.309)	0.091 (0.297)	Often share on FB	-0.086 (0.129)	0.035 (0.131)	Religious	0.056 (0.176)	0.291* (0.171)
Age	-0.022 (0.030)	-0.022 (0.030)	Reason to share: interest	0.063 (0.312)	0.211 (0.316)	Very religious	0.077 (0.275)	-0.149 (0.273)
Age squared	0.000 (0.000)	0.000 (0.000)	Reason to share: influence	0.411 (0.298)	0.806*** (0.303)	Religion: catholic	0.263** (0.125)	0.218* (0.124)
Male	0.017 (0.126)	-0.067 (0.121)	Reason to share: image	-0.571** (0.272)	-0.177 (0.277)	Religion: muslim	0.030 (0.288)	-0.015 (0.289)
Born in France	-0.173 (0.321)	0.337 (0.326)	Reason to share: reciprocity	-0.150 (0.285)	-0.405 (0.293)	Feel closer to European government	0.135 (0.188)	0.117 (0.196)
Live in village	-0.219 (0.166)	-0.309* (0.162)	Source news: TV	0.276 (0.214)	0.067 (0.204)	Feel closer to French government	0.023 (0.119)	0.126 (0.120)
Live in city	-0.074 (0.147)	-0.314** (0.147)	Source news: Internet	0.008 (0.221)	-0.198 (0.211)	Negative Image EU (1-5)	0.012 (0.056)	0.001 (0.057)
Father born in France	0.052 (0.205)	-0.062 (0.210)	Source news: Radio	-0.095 (0.303)	0.177 (0.288)	Voted Le Pen, 1st round 2017	-0.237 (0.202)	-0.396* (0.210)
Mother born in France	0.443* (0.230)	0.279 (0.223)	Source news: newspapers	0.204 (0.279)	-0.031 (0.276)	Voted Macron, 1st round 2017	-0.135 (0.169)	0.094 (0.165)
Married	0.137 (0.137)	0.281** (0.141)	Self-reported altruism	0.306 (0.778)	-0.031 (0.771)	Voted Le Pen, 2nd round 2017	0.062 (0.196)	0.150 (0.203)
Single	-0.198 (0.164)	0.044 (0.161)	Self-reported reciprocity	0.141 (0.355)	-0.420 (0.353)	Registered to vote in 2019	0.161 (0.168)	-0.094 (0.164)
Income (1-10)	0.002 (0.029)	0.004 (0.030)	Self-reported image importance	-0.322 (0.254)	-0.190 (0.258)	Regional dummies	✓	✓
Children	-0.178 (0.146)	0.045 (0.142)	Gave money to charity	-0.024 (0.346)	-0.140 (0.345)	Observations	2,133	2,133
						p-value: joint significance, per treatment	0.744	0.165
						p-value: joint significance both treatments	0.260	

Note: The table presents the results of a multinomial logit regression, in which treatment status is regressed on all pre-treatment characteristics. At the end of the table, we report p-values from the tests of joint significance of all pre-treatment characteristics in explaining: (1) Alt-Facts vs. Imposed Fact Check, (2) Alt-Facts vs. Voluntary Fact Check, (3) any treatment status. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A2: Comparison of our sample with Facebook users

	Our sample		Facebook users
	mean	std.dev.	mean
Age	44.36	15.08	34
Male	0.50	0.50	0.49
Married	0.39	0.49	0.28
In a civil partnership	0.25	0.43	0.26
Single	0.24	0.43	0.32
Higher education	0.28	0.45	0.36
High school diploma	0.15	0.35	0.18
Daily use of Facebook	0.88	0.32	0.77
Number of friends	214	483	338

Note: These descriptive statistics of the Facebook users come from a variety of sources: gender, marital status and education for French users are obtained from: <https://blog.digimind.com/fr/agences/facebook-chiffres-essentiels> (accessed on December 24, 2020), age distribution is obtained from Digital 2020 global report and refers to the whole world, rather than France. Information of daily users is from <https://www.blogdumoderateur.com/chiffres-facebook/> (accessed on December 24, 2020). Average number of friends on Facebook is from <https://www.brandwatch.com/blog/facebook-statistics/> (accessed on December 24, 2020) and is a world average.

Table A3: Average treatment effect on sharing alt-facts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Intent to Share Alt-Facts on Facebook				Action of Sharing Alt-Facts on Facebook			
Imposed Fact-Check	-0.045*** (0.016)	-0.045*** (0.016)	-0.049*** (0.017)	-0.050*** (0.017)	-0.020*** (0.006)	-0.020*** (0.006)	-0.021*** (0.006)	-0.021*** (0.007)
Voluntary Fact-Check	-0.038** (0.016)	-0.038** (0.016)	-0.040** (0.017)	-0.052*** (0.017)	-0.021*** (0.005)	-0.021*** (0.005)	-0.024*** (0.006)	-0.026*** (0.006)
Male		0.040*** (0.014)	0.047*** (0.015)	0.029* (0.016)		0.011** (0.005)	0.014*** (0.005)	0.009* (0.005)
Low Education		0.011 (0.019)	0.094** (0.046)	0.106** (0.047)		0.010* (0.006)	0.017 (0.014)	0.018 (0.015)
Middle education		0.050* (0.027)	0.052* (0.028)	0.055* (0.028)		0.015* (0.008)	0.012 (0.009)	0.014 (0.010)
Age			-0.001 (0.003)	-0.003 (0.004)			0.000 (0.001)	-0.000 (0.001)
Age squared			-0.000 (0.000)	0.000 (0.000)			-0.000 (0.000)	0.000 (0.000)
Income (1-10)			-0.004 (0.004)	-0.001 (0.004)			-0.003*** (0.001)	-0.003** (0.001)
Religious			0.105*** (0.021)	0.036 (0.023)			0.029*** (0.007)	0.014* (0.008)
Frequency of FB use			-0.008 (0.010)	-0.005 (0.011)			-0.005 (0.004)	-0.005 (0.004)
Often share on FB			0.060*** (0.015)	0.036** (0.016)			0.016*** (0.005)	0.009* (0.005)
Log (FB friends+1)			0.001 (0.006)	-0.006 (0.006)			0.002 (0.002)	0.000 (0.002)
Voted Le Pen, 2nd round 2017			0.175*** (0.019)	0.144*** (0.028)			0.045*** (0.007)	0.030*** (0.008)
Negative Image EU (1-5)			-0.004 (0.008)	-0.001 (0.008)			-0.001 (0.003)	-0.001 (0.003)
Self-reported altruism				0.151 (0.101)				0.060 (0.038)
Self-reported reciprocity				-0.027 (0.044)				-0.007 (0.020)
Self-reported image importance				-0.040 (0.032)				-0.009 (0.013)
Reason to share: interest				-0.041 (0.037)				-0.017 (0.015)
Reason to share: influence				0.163*** (0.040)				0.042*** (0.014)
Reason to share: image				0.001 (0.033)				0.004 (0.012)
Reason to share: reciprocity				0.059* (0.033)				0.017 (0.012)
Observations	2,537	2,537	2,265	2,078	2,537	2,537	2,265	2,078
R-squared	0.004	0.010	0.110	0.166	0.008	0.010	0.073	0.101
Mean DV, Alt-Facts Treatment	0.147	0.147	0.155	0.161	0.0462	0.0462	0.0492	0.0509
p-value, Imposed=Voluntary	0.635	0.634	0.530	0.899	0.812	0.812	0.579	0.379
Demographic controls			✓	✓			✓	✓
All pretreatment characteristics				✓				✓

Note: The comparison group is Alt-Facts treatment. The set of unreported demographic controls is as follows: family status (dummies for married and single), number of children, size of the place of living (dummies for village and town), religion (dummies for catholic, muslim, and no religion), and dummies for each of the 9 levels of education. The list of all pretreatment characteristics is given in Table A1. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Average treatment effect on sharing fact-checking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Intent to Share Fact-Check on Facebook				Action of Sharing Fact-Check on Facebook			
Voluntary Fact-Check	-0.028*	-0.028*	-0.041**	-0.039**	-0.001	-0.001	-0.001	0.004
	(0.016)	(0.016)	(0.017)	(0.018)	(0.006)	(0.006)	(0.006)	(0.007)
Male		0.046***	0.048***	0.032		0.013**	0.014*	0.007
		(0.017)	(0.018)	(0.020)		(0.006)	(0.007)	(0.007)
Low Education		-0.136***	0.012	0.010		-0.007	0.018	0.026
		(0.029)	(0.060)	(0.064)		(0.009)	(0.015)	(0.025)
Middle education		-0.058	-0.011	-0.014		-0.010	0.000	-0.000
		(0.038)	(0.041)	(0.044)		(0.009)	(0.010)	(0.011)
Age			0.007**	0.005			0.001	0.001
			(0.004)	(0.005)			(0.001)	(0.002)
Age squared			-0.000**	-0.000			-0.000	-0.000
			(0.000)	(0.000)			(0.000)	(0.000)
Income (1-10)			0.002	-0.001			-0.002	-0.003
			(0.005)	(0.005)			(0.002)	(0.002)
Religious			0.063***	0.018			0.013	0.004
			(0.024)	(0.028)			(0.009)	(0.010)
Frequency of FB use			-0.012	-0.023*			-0.007	-0.009*
			(0.012)	(0.014)			(0.005)	(0.005)
Often share on FB			0.072***	0.044**			0.013*	0.005
			(0.019)	(0.021)			(0.007)	(0.008)
Log (FB friends+1)			0.007	0.001			0.004*	0.001
			(0.007)	(0.008)			(0.002)	(0.003)
Voted Le Pen, 2nd round 2017			-0.008	0.029			0.001	0.008
			(0.019)	(0.029)			(0.007)	(0.009)
Negative Image EU (1-5)			-0.049***	-0.036***			-0.014***	-0.010***
			(0.010)	(0.010)			(0.003)	(0.003)
Self-reported altruism				0.257*				0.111**
				(0.132)				(0.048)
Self-reported reciprocity				-0.002				0.007
				(0.056)				(0.019)
Self-reported image importance				-0.018				0.014
				(0.041)				(0.014)
Reason to share: interest				0.091*				0.005
				(0.048)				(0.016)
Reason to share: influence				0.157***				0.050***
				(0.050)				(0.019)
Reason to share: image				-0.033				-0.032*
				(0.044)				(0.017)
Reason to share: reciprocity				0.005				0.023
				(0.047)				(0.018)
Observations	1,692	1,692	1,517	1,388	1,692	1,692	1,517	1,388
R-squared	0.002	0.038	0.100	0.131	0.000	0.004	0.038	0.073
Mean DV, Imposed Fact-Check	0.143	0.143	0.158	0.160	0.0315	0.0315	0.0337	0.0320
Demographic controls			✓	✓			✓	✓
All pretreatment characteristics				✓				✓

Note: The comparison group is Imposed Fact-Check treatment. The set of unreported demographic controls is as follows: family status (dummies for married and single), number of children, size of the place of living (dummies for village and town), religion (dummies for catholic, muslim, and no religion), and dummies for each of the 9 levels of education. The list of all pretreatment characteristics is given in Table A1. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Predictors of sharing alt-facts and sharing fact-check:
variables selected by adaptive LASSO

Dep. var.:	(1)	(2)	(3)	(4)
	Sharing Intent Alt-Facts	Sharing Action Alt-Facts	Sharing Intent Fact-Check	Sharing Action Fact-Check
Sample:	Alt-Facts Treatment		Imposed Fact-Check Treatment	
Voted Le Pen, 2nd round 2017	0.197*** (0.051)	0.073*** (0.012)		
Voted Le Pen, 1st round 2017	0.057 (0.053)			
Feel closer to European gov.	-0.080*** (0.024)	-0.039*** (0.008)	0.105** (0.050)	
Feel closer to French gov.				-0.013* (0.007)
Negative Image EU (1-5)			-0.026* (0.014)	-0.009** (0.004)
Education level (1-9)			0.016*** (0.005)	
Male	0.053** (0.024)	0.027*** (0.009)	0.048** (0.024)	0.014** (0.007)
Frequency of FB use			-0.031* (0.017)	-0.011** (0.005)
Reason to share: interest			0.156** (0.062)	
Reason to share: influence	0.257*** (0.057)	0.054** (0.021)	0.081 (0.061)	0.037* (0.019)
Reason to share: reciprocity			0.093 (0.059)	0.023 (0.015)
Source news: Internet			-0.034 (0.026)	
Self-reported reciprocity			0.086 (0.067)	0.046*** (0.018)
Gave money to charity			0.130** (0.054)	0.039*** (0.015)
Gave money to homeless			0.048 (0.059)	
Donated blood	0.076* (0.041)	0.026* (0.015)	0.062 (0.043)	
Log (FB friends+1)	-0.014 (0.009)			
Religious	0.145*** (0.036)	0.056*** (0.016)		
Father born in France				-0.015 (0.014)
Married				0.014* (0.008)
Income not reported	-0.102*** (0.029)			
Observations	778	790	773	781
R-squared	0.196	0.136	0.133	0.082

Note: OLS regression results reported. The explanatory variables in each of these regressions were selected by LASSO from all 37 pretreatment characteristics. Standard errors are corrected for heteroscedasticity. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Table A6: Actual compared to predicted sharing for viewers and nonviewers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Difference b/w actual and predicted: Intent to Share Alt-Facts on Facebook				Difference b/w actual and predicted: Action of Sharing Alt-Facts on Facebook			
Imposed Fact-Check	-0.046*** (0.016)	-0.045*** (0.016)	-0.052*** (0.017)	-0.051*** (0.018)	-0.021*** (0.006)	-0.021*** (0.006)	-0.022*** (0.007)	-0.021*** (0.007)
Voluntary Fact-Check: Viewer	-0.025 (0.023)	-0.028 (0.023)	-0.028 (0.024)	-0.025 (0.024)	-0.023*** (0.008)	-0.022*** (0.008)	-0.022*** (0.008)	-0.020** (0.008)
Voluntary Fact-Check: Nonviewer	-0.066*** (0.017)	-0.063*** (0.017)	-0.074*** (0.018)	-0.072*** (0.019)	-0.026*** (0.006)	-0.027*** (0.006)	-0.032*** (0.007)	-0.030*** (0.007)
Male		-0.019 (0.014)	-0.014 (0.016)	-0.021 (0.016)		-0.017*** (0.005)	-0.013** (0.005)	-0.016*** (0.006)
Low Education		-0.029 (0.019)	0.094* (0.048)	0.111** (0.048)		-0.001 (0.006)	0.016 (0.014)	0.017 (0.015)
Middle education		0.027 (0.026)	0.047* (0.029)	0.058** (0.029)		0.010 (0.009)	0.011 (0.010)	0.014 (0.010)
Age			-0.003 (0.003)	-0.004 (0.004)			0.000 (0.001)	-0.000 (0.001)
Age squared			0.000 (0.000)	0.000 (0.000)			-0.000 (0.000)	0.000 (0.000)
Income (1-10)			-0.003 (0.004)	-0.002 (0.004)			-0.003** (0.001)	-0.003* (0.001)
Religious			-0.047** (0.022)	-0.097*** (0.023)			-0.022*** (0.008)	-0.035*** (0.008)
Frequency of FB use			-0.009 (0.011)	-0.005 (0.011)			-0.006 (0.004)	-0.005 (0.004)
Often share on FB			0.030* (0.015)	0.035** (0.016)			0.010* (0.005)	0.010* (0.005)
Log (FB friends+1)			0.000 (0.006)	0.001 (0.006)			0.000 (0.002)	0.000 (0.002)
Voted Le Pen, 2nd round 2017			-0.073*** (0.020)	-0.057** (0.028)			-0.032*** (0.007)	-0.045*** (0.008)
Negative Image EU (1-5)			-0.005 (0.008)	-0.001 (0.008)			-0.002 (0.003)	-0.001 (0.003)
Self-reported altruism				-0.103 (0.101)				0.019 (0.038)
Self-reported reciprocity				-0.016 (0.044)				-0.007 (0.020)
Self-reported image importance				-0.048 (0.032)				-0.009 (0.013)
Reason to share: interest				-0.046 (0.038)				-0.018 (0.015)
Reason to share: influence				-0.090** (0.040)				-0.013 (0.014)
Reason to share: image				0.008 (0.033)				0.004 (0.012)
Reason to share: reciprocity				0.055 (0.034)				0.017 (0.012)
Observations	2,219	2,219	2,051	2,018	2,349	2,349	2,112	2,078
R-squared	0.007	0.011	0.038	0.069	0.010	0.016	0.046	0.063
Mean Dep. var., Alt-Facts T	-0.0014	-0.0014	0.0036	0.0023	0.0006	0.0006	0.0027	0.0019
Mean predicted Alt-Facts T	0.150	0.150	0.156	0.159	0.0474	0.0474	0.0483	0.0490
Mean Predict. Imposed Fact-Check T	0.158	0.158	0.159	0.159	0.0467	0.0467	0.0479	0.0480
Mean Predict., Voluntary T, Viewer	0.188	0.188	0.190	0.190	0.0589	0.0589	0.0592	0.0592
Mean Predict., Voluntary T, Nonviewer	0.136	0.136	0.144	0.144	0.0432	0.0432	0.0442	0.0442
Demographic controls			✓	✓			✓	✓
All pretreatment characteristics				✓				✓

Note: The comparison group is Alt-Facts treatment. The set of unreported demographic controls is as follows: family status (dummies for married and single), number of children, size of the place of living (dummies for village and town), religion (dummies for catholic, muslim, and no religion), and dummies for each of the 9 levels of education. The list of all pretreatment characteristics is given in Table A1. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Predictors of sharing behavior for alt-facts only, fact-check only, and both alt-facts and fact-check: variables selected by adaptive LASSO

Dep. var.:	(1) Sharing Intent Alt-Facts only	(2) Sharing Action Alt-Facts only	(3) Sharing Intent Fact-Check only	(4) Sharing Action Fact-Check only	(5) Sharing Intent Both	(6) Sharing Action Both
Sample: Treatment	Alt-Facts		Imposed Fact-Check		Imposed Fact-Check	
Voted Le Pen, 2nd round 2017	0.197*** (0.051)	0.073*** (0.012)	-0.054*** (0.019)	-0.010* (0.006)	0.064*** (0.021)	0.003 (0.002)
Voted Le Pen, 1st round 2017	0.057 (0.053)					0.003 (0.002)
Feel closer to European gov.	-0.080*** (0.024)	-0.039*** (0.008)	0.094** (0.045)			
Feel closer to French gov.				-0.014** (0.007)		
Negative Image EU (1-5)			-0.018 (0.012)	-0.006* (0.003)	-0.015 (0.010)	-0.002 (0.002)
Age			0.002** (0.001)		-0.000 (0.000)	
Single					0.017 (0.019)	
Education level (1-9)			0.011** (0.005)		0.006** (0.003)	
Male	0.053** (0.024)	0.027*** (0.009)	0.040* (0.021)	0.015** (0.007)		
Father born in France			-0.037 (0.036)	-0.015 (0.014)		
Frequency of FB use			-0.019 (0.015)	-0.011** (0.005)		
Log (FB friends+1)	-0.014 (0.009)					
Reason to share: interest			0.163*** (0.048)			
Reason to share: influence	0.257*** (0.057)	0.054** (0.021)		0.030 (0.018)	0.107*** (0.036)	0.008*** (0.002)
Reason to share: reciprocity				0.022 (0.015)		
Source news: TV			-0.062** (0.031)			
Source news: Internet			-0.066** (0.033)			0.003 (0.002)
Self-reported reciprocity			0.123** (0.059)	0.047*** (0.018)		
Gave money to charity			0.110** (0.052)	0.032** (0.014)		
Worked for charity			-0.082* (0.049)			
Religion: catholic			-0.032 (0.023)	-0.011 (0.007)		
Donated blood	0.076* (0.041)	0.026* (0.015)				
Religious	0.145*** (0.036)	0.056*** (0.016)			0.023 (0.019)	
Married				0.017** (0.008)		-0.002* (0.001)
Live in city					-0.035** (0.014)	-0.003* (0.002)
Often share on FB					0.029** (0.012)	
Self-reported altruism					0.150*** (0.049)	0.013** (0.006)
Income not reported	-0.102*** (0.029)			0.003 (0.031)		
Observations	778	790	773	781	771	678
R-squared	0.196	0.136	0.117	0.083	0.117	0.073

Note: OLS regression results reported. The explanatory variables in each of these regressions were selected by LASSO from all 37 pretreatment characteristics. Standard errors are corrected for heteroscedasticity. * p \leq 0.05, ** p \leq 0.01, *** p \leq 0.001.

Table A8: Testing for experimenter demand effects: heterogeneity by social desirability score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Intent to Share Alt-Facts on Facebook				Action of Sharing Alt-Facts on Facebook			
Imposed Fact-Check	-0.041** (0.016)	-0.041** (0.016)	-0.046*** (0.016)	-0.050*** (0.017)	-0.019*** (0.006)	-0.019*** (0.006)	-0.020*** (0.006)	-0.021*** (0.007)
Voluntary Fact-Check	-0.034** (0.016)	-0.034** (0.016)	-0.037** (0.016)	-0.052*** (0.017)	-0.020*** (0.005)	-0.020*** (0.005)	-0.023*** (0.006)	-0.026*** (0.006)
Imposed Fact-Check × × Social Desirability Score	-0.022 (0.015)	-0.022 (0.015)	-0.015 (0.016)	-0.011 (0.016)	-0.009 (0.006)	-0.009 (0.006)	-0.007 (0.006)	-0.007 (0.006)
Voluntary Fact-Check × × Social Desirability Score	-0.007 (0.015)	-0.009 (0.015)	-0.001 (0.015)	0.007 (0.016)	-0.005 (0.005)	-0.006 (0.005)	-0.004 (0.005)	-0.002 (0.006)
Social Desirability Score	0.050*** (0.011)	0.050*** (0.011)	0.032*** (0.012)	–	0.015*** (0.004)	0.015*** (0.004)	0.011** (0.004)	–
Observations	2,537	2,537	2,265	2,078	2,537	2,537	2,265	2,078
R-squared	0.026	0.030	0.119	0.167	0.020	0.022	0.078	0.101
Strata controls		✓	✓	✓		✓	✓	✓
Demographic controls & Facebook use			✓	✓			✓	✓
All pretreatment characteristics				✓				✓

Note: The comparison group is Alt-Facts treatment. Controls are the same as in the respective columns of Table 2. Coefficients for the social desirability score are not reported in columns (4) and (8) as the social desirability score is a linear combination of three pretreatment variables. Social Desirability Score has mean of zero and SD of 1.16. Robust standard errors are in parentheses. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

B Questions on sharing intention and actual sharing in our experiment and in related studies

B.1 Questions on sharing intention and actual sharing in our experiment

French original (English translation is in Figure 1 of the main text):

Screen 1: measurement of “Sharing Intention” variable

Voulez vous partager cet article "RN : L'Union Européenne contrôle 87% de nos lois et souhaite ouvrir la porte à 50 millions d'immigrés", contenant les déclarations de Marine Le Pen et Jordan Bardella sur votre mur Facebook ?

- Oui
- Non
- Ne souhaite pas répondre

Screen 2: measurement of “Sharing Action” variable

Vous avez la possibilité de partager l'article que vous venez de lire sur votre mur Facebook en cliquant sur ce bouton de partage (si vous n'êtes pas déjà connecté sur Facebook et que vous souhaitez partager le contenu, vous pouvez soit vous connecter maintenant sur une page séparée, soit vous serez amené à vous connecter lorsque vous cliquerez sur le bouton de partage) :

 PARTAGER

B.2 Questions about sharing alt-facts used in related studies

In this section, we list the formulation of the question about sharing fake news used in the related literature (cited and described in the main text). All of them are hypothetical questions about self-reported willingness to share. In contrast, as one of the two main outcomes in our study, we use clicking on the Facebook share button. In addition, there are two important differences in our self-reported intention to share alt-facts variable compared to the formulations adopted by other studies. First, the formulations of the intent to share question in these papers is much more hypothetical than in our case. Second, these questions do not specify the specific platform on which sharing takes place.

List of questions about sharing alt-facts in other studies:

- Fazio (2020):

“How likely would you be to share this story online?”

- Pennycook et al. (2020a):

“We are interested in whether you would consider sharing these stories on social media (such as Facebook or Twitter).”

- Pennycook et al. (2020b):

“Would you consider sharing this story online (for example, through Facebook or Twitter?)”

- Yaqub et al. (2020):

“Would you share this headline with your friends on social media?”

- Altay et al. (2020):

“How likely would you be to share this story online (for example, through Facebook or Twitter)?”

C Treatments

C.1 Screenshot of the Alt-Facts as they appear on respondents' screens

Page 1 of 2 (See Subsection C.3 for English translation):

02/08/2019

Qualtrics Survey Software

RN : L'Union Européenne contrôle 87% de nos lois et souhaite ouvrir la porte à 50 millions d'immigrés



Dans une interview le 25 juin 2018 sur franceinfo, Marine Le Pen pointait du doigt la volonté de l'Union Européenne d'ouvrir la porte aux immigrés :

« L'Union européenne souhaite de l'immigration. Elle l'a dit à plusieurs reprises par l'intermédiaire de la bouche, d'ailleurs, de beaucoup de ses commissaires européens. Ils sont même allés jusqu'à dire 50 millions d'immigrés d'ici 2050.»

0:00 / 0:36

Jordan Bardella, la tête de liste du Rassemblement National pour les élections européennes, pointait du doigt, le 4 juin 2018 sur Sud Radio, un autre type de risque

« Il nous faut retrouver la maîtrise de nos budgets, retrouver la maîtrise de nos lois. Je rappelle que 87 % de nos lois, des lois qui sont votées, viennent de directives européennes. »

The screenshot shows the Sud Radio website interface. At the top, there is a navigation bar with the Sud Radio logo, social media icons (Facebook, Twitter), a search icon, and a 'ÉCOUTER LE DIRECT' button. Below the navigation bar, there is a header with 'FIL INFO' and a main headline: '10:45 >> Des femmes catholiques dénoncent une Eglise qui se rigidifie et "tombe en ruine"'. Below the headline, there is a sub-headline: 'Par Benjamin Jeanjean / Lundi 4 juin 2018 à 08:34'. The main content area features a video player showing Jordan Bardella speaking into a microphone. To the right of the video player, there is a section titled 'LES ÉDITOS SUD RADIO' with a sub-headline: 'EXCLU : Ce que l'on sait de la polémique du mystérieux candidat du...'. Below this, there is a list of editors: 'Véronique Jacquier' with the sub-headline 'Attentat déjoué, Black Bloc : nos policiers sont-ils en danger ?' and 'Yoaine de La Bigne'.

C.2 Screenshot of the Fact-Checking as it appear on respondents' screens

Page 1 of 3 (See the fact-checking part of Subsection C.4 for English translation):

02/08/2019

Qualtrics Survey Software

Plusieurs articles de fact checking ont jugé les déclarations que vous avez vues comme fausses, voici un de ces articles.

Intox: Non L'Union Européenne ne contrôle pas 87% de nos lois et n'organise pas un afflux de 50 millions d'immigrés !



« Il nous faut retrouver la maîtrise de nos budgets, retrouver la maîtrise de nos lois. Je rappelle que 87 % de nos lois, des lois qui sont votées, viennent de directives européennes.» Jordan Bardella, 4 juin 2018

Le mythe que plus de 80% de nos lois sont imposées par l'Europe trouve ses origines dans une déclaration de Jacques Delors, le 14 avril 1988, "L'Europe, ce sera en 1992, 80% de la législation économique, financière et fiscale qui sera d'origine européenne". Sa prédiction est loin d'être vérifiée.

<https://pseparis.eu.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrintPreview>

15/25

En 2014 l'institut Notre Europe montrait que c'est plutôt 20% de la législation française qui venait directement ou indirectement de décisions prises au niveau européen. Dans [un papier de 2009 intitulé "La législation nationale d'origine communautaire : briser le mythe des 80%"](#), Bertoncini montre que les actes communautaires ne représentent que 36,2% des textes français.

Une étude plus récente – "Les lois françaises sont-elles écrites à Bruxelles ?" – [publiée dans la Revue de l'Union européenne en 2012](#) la proportion du nombre total de lois européennes atteint 13,3% en 2006. Dans [un papier écrit pour la London School of Economics and Political Science](#), Annette Elisabeth Toeller de l'Université de Hagen: "La conclusion la plus frappante est que la plupart de ces études ont montré un taux relativement bas d'eupéanisation de la législation européenne : [15,5% au Royaume-Uni](#), [14% au Danemark](#), 10,6% pour l'Autriche, entre 3 et 27% pour la France, entre 1 et [24% pour la Finlande](#), mais de [39,1% en Allemagne](#)".

Même si le calcul est compliqué, et que les différentes études ne s'accordent pas sur un seul chiffre, toutes ces études montrent qu'on est largement en dessous de 50% des lois venant des directives européennes et donc très loin des 87% cités par Jordan Bardella.

[Source fact checking 1](#)

[Source fact checking 2](#)

« L'Union européenne souhaite de l'immigration. Elle l'a dit à plusieurs reprises par l'intermédiaire de la bouche, d'ailleurs, de beaucoup de ses commissaires européens. Ils sont même allés jusqu'à dire 50 millions d'immigrés d'ici 2050. » Marine Le Pen, 25 juin 2018

Le chiffre de 50 million apparait dans une déclaration le 8 juin du commissaire européen aux Affaires intérieures, Dimitris Avramopoulos: *«On sait tous que l'Europe est un continent qui vieillit. Sans immigration, la population active européenne déclinera de plus de 20 millions dans les quinze années à venir. Et d'ici 2060, la population active diminuera de 10%, c'est-à-dire de 50 millions.»*

Même si le constat du vieillissement de la population est souvent fait, la solution proposée est rarement celle d'encourager l'immigration, ou si c'est le cas, seulement l'immigration choisie comme le propose [ce rapport](#).

Un récent document de la commission auquel le RN se réfère indique «Le fait est que, d'ici 2050, en l'absence, peu probable, d'immigration et à taux d'activité constant, la population active de l'UE diminuera d'environ 68 millions de travailleurs. Etant donné que tous les immigrés ne rejoignent pas la population active, il faudrait un gain net de 100 millions de personnes environ pour combler le déficit. **Objectivement, un afflux net aussi important au cours des quarante prochaines années n'est ni probable, ni nécessairement souhaitable.** »

Enfin et surtout, la Commission européenne n'a pas compétence en matière d'admission au séjour dans les Etats membres. Ce qui est écrit très clairement, également, dans cette [communication de la Commission au Parlement](#) : «Les Etats membres conserveront la compétence exclusive lorsqu'il s'agira de décider du volume des admissions de ressortissants de pays tiers qui émigrent pour chercher du travail.»

[Source fact checking](#)

C.3 Translation of the Text of the Alt-Facts Treatment

We are going to show you a set of articles on the European Union. Please read them carefully.

Article 1 RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants.

In an interview on the 25th of June 2018 Marine Le Pen accused the European Union of wanting to open the doors to immigrants:

“The European Union wants immigration. It has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 millions immigrants by 2050.”

On the 4th of June 2018 on Sud Radio, Jordan Bardella, the lead candidate for the National Rally at the European Elections, pointed to another risk

"We need to regain control over our budget, regain control of our laws. I remind you that 87% of our laws, of the laws that are voted, comes from European directives"

Q29 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella on your Facebook page?

1. Yes
2. No
3. I prefer not to answer this question

You can share the article that you have just read on your Facebook page by clicking on this sharing button (if you are not already connected on Facebook and you want to share the content you can either connect yourself on a separate page or you will be redirected to Facebook login page after clicking on the button)

Q30 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella with 100 other participants that will take this survey after you?

1. Yes
2. No
3. I prefer not to answer this question

C.4 Translation of the text of the Imposed Fact-Check Treatment

We are going to show you a set of articles on the European Union. Please read them carefully.

Article 1 RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants.

In an interview on the 25th of June 2018 Marine Le Pen accused the European Union of wanting to open the doors to immigrants:

“The European Union wants immigration. It has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 millions immigrants by 2050.”

On the 4th of June 2018 on Sud Radio, Jordan Bardella, the lead candidate for the National Rally at the European Elections, pointed to another risk

"We need to regain control over our budget, regain control of our laws. I remind you that 87% of our laws, of the laws that are voted, comes from European directives"

Several fact-checking articles have shown these claims are false.

Intox: the European Union does not control 87% of our laws and is not organizing an influx of 50 millions immigrants!

“We need to regain control over our budget, regain control of our laws. I remind you that 87% of our laws, of the laws that are voted, comes from European directives.”

The myth of the European Union imposing 80% of laws originated in a Jacques Delors' statement made on the 14th of April 1988: “By 1992, 80% of the economic, financial and fiscal legislation will be of European origin". His prediction is far from being met.

In 2014, the Notre Europe institute showed that 20% of French laws came directly or indirectly from a decision taken at the European level. In a 2009 paper entitled “The national legislation of European origin: breaking the myth” Bertoincini shows that European acts represent only 36.2% of the French laws.

A more recent study—“Are French laws written in Bruxelles?”—published in the *Revue de l’Union Européenne* in 2012, the proportion of laws with European origins reaches 13.3%. In a paper write for the LSE, Annette Elisabeth Toeller of the University of Hagen: “The most striking conclusion is that most of these studies show a relatively low rate of europenisation of European laws: 15.5% in the UK, 14% in Denmark, 10.6% in Austria, between 3 and 27% in France, between 1 and 24% in Finland, but 39.1% in Germany.”

Even if estimating this proportion precisely is not an easy task and these studies do not agree on the exact number, all of them show that we are well below 50% and therefore far from the 87% referred to by Jordan Bardella.

"The European Union wants immigration. It has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 millions immigrants by 2050."

The 50 millions figure appeared in a declaration made by the European commissioner to Interior Affairs, Dimitris Avramopoulos on the 8th of June: “We all know that Europe is an ageing continent. Without immigration, the active population will fall by 20 million individuals in the next 15 years. And by 2060, the active population will shrink by 10%, meaning by 50 millions.”

Even if the problem of an ageing population is often mentioned, the proposed solution is rarely to promote immigration and if this is the case, only a regulated type of immigration as the one proposed in this document.

A recent document by the Commission mentioned by the RN says: “The fact is that by 2050, in the improbable absence of immigration and with a constant rate of participation to the labour market, the active population in the EU will decrease by 68 millions workers. Given that not all of the immigrants join the active workforce, we will need a net increase of around 100 million people in order to close the deficit.” Objectively, an influx of this magnitude in the next 40 years is neither likely and nor necessarily desirable.”

Finally and most importantly, the European Commission does not have the authority over admitting residents into European member states. This is clearly stated in this communication by the Commission to the Parliament: “The member States will maintain their exclusive right on matters related to the numbers of admissions of non-EU migrants in search of work.”

Q29 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella on your Facebook page?

1. Yes
2. No
3. I prefer not to answer this question

You can share the article that you have just read on your Facebook page by clicking on this sharing button (if you are not already connected on Facebook and you want to share the content you can either connect yourself on a separate page or you will be redirected to Facebook login page after clicking on the button)

Q30 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella with 100 other participants that will take this survey after you?

1. Yes
2. No
3. I prefer not to answer this question

Q31 Do you want to share the fact-checking article on your Facebook page?

1. Yes
2. No
3. I prefer not to answer this question

You can share the article that you have just read on your Facebook page by clicking on this sharing button (if you are not already connected on Facebook and you want to share the content you can either connect yourself on a separate page or you will be redirected to Facebook login page after clicking on the button).

C.5 Translation of the text of the Voluntary Fact-Check Treatment

We are going to show you a set of articles on the European Union.
Please read them carefully.

Article 1 RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants.

In an interview on the 25th of June 2018 Marine Le Pen accused the European Union of wanting to open the doors to immigrants:

“The European Union wants immigration. It has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 millions immigrants by 2050.”

On the 4th of June 2018 on Sud Radio, Jordan Bardella, the lead candidate for the National Rally at the European Elections, pointed to another risk

"We need to regain control over our budget, regain control of our laws. I remind you that 87% of our laws, of the laws that are voted, comes from European directives"

Q28 Do you want to read some fact-checking articles concerning the claims that you have just read?

1. Yes
2. No
3. I prefer not to answer this question

IF THE RESPONDENTS ANSWERED YES, SHOW THE REST.

IF THE RESPONDENT ANSWERED NO, SKIP TO QUESTIONS Q29 AND Q30. DO NOT ASK QUESTION Q31.

Several fact-checking articles have shown these claims are false.

Intox: the European Union does not control 87% of our laws and is not organizing an influx of 50 millions immigrants!

“We need to regain control over our budget, regain control of our laws. I remind you that 87% of our laws, of the laws that are voted, comes from European directives.”

The myth of the European Union imposing 80% of laws originated in a Jacques Delors' statement made on the 14th of April 1988: "By 1992, 80% of the economic, financial and fiscal legislation will be of European origin". His prediction is far from being met.

In 2014, the Notre Europe institute showed that 20% of French laws came directly or indirectly from a decision taken at the European level. In a 2009 paper entitled "The national legislation of European origin: breaking the myth" Bertonecini shows that European acts represent only 36.2% of the French laws.

A more recent study—"Are French laws written in Bruxelles?"—published in the *Revue de l'Union Européenne* in 2012, the proportion of laws with European origins reaches 13.3%. In a paper written for the LSE, Annette Elisabeth Toeller of the University of Hagen: "The most striking conclusion is that most of these studies show a relatively low rate of europeanisation of European laws: 15.5% in the UK, 14% in Denmark, 10.6% in Austria, between 3 and 27% in France, between 1 and 24% in Finland, but 39.1% in Germany."

Even if estimating this proportion precisely is not an easy task and these studies do not agree on the exact number, all of them show that we are well below 50% and therefore far from the 87% referred to by Jordan Bardella.

"The European Union wants immigration. It has said this multiple times, through the voices, among others, of multiple European Commissioners. They even went as far as saying 50 millions immigrants by 2050."

The 50 millions figure appeared in a declaration made by the European commissioner to Interior Affairs, Dimitris Avramopoulos on the 8th of June: "We all know that Europe is an ageing continent. Without immigration, the active population will fall by 20 million individuals in the next 15 years. And by 2060, the active population will shrink by 10%, meaning by 50 millions."

Even if the problem of an ageing population is often mentioned, the proposed solution is rarely to promote immigration and if this is the case, only a regulated type of immigration as the one proposed in this document.

A recent document by the Commission mentioned by the RN says: "The fact is that by 2050, in the improbable absence of immigration and with a constant rate of participation to the labour market, the active population in the EU will decrease by 68 millions workers. Given that not all of the immigrants join the active workforce, we will need a net increase of around 100 million people in order to close the deficit." Objectively, an influx of this magnitude in the next 40

years is neither likely and nor necessarily desirable.”

Finally and most importantly, the European Commission does not have the authority over admitting residents into European member states. This is clearly stated in this communication by the Commission to the Parliament: “The member States will maintain their exclusive right on matters related to the numbers of admissions of non-EU migrants in search of work.”

Q29 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella on your Facebook page?

1. Yes
2. No
3. I prefer not to answer this question

You can share the article that you have just read on your Facebook page by clicking on this sharing button (if you are not already connected on Facebook and you want to share the content you can either connect yourself on a separate page or you will be redirected to Facebook login page after clicking on the button)

Q30 Do you want to share the article “RN: the European Union controls 87% of our laws and aims at opening the door to 50 millions immigrants” containing the claims made by Marine Le Pen and Jordan Bordella with 100 other participants that will take this survey after you?

1. Yes
2. No
3. I prefer not to answer this question

Q31 Do you want to share the fact-checking article on your Facebook page?

1. Yes
2. No
3. I prefer not to answer this question

You can share the article that you have just read on your Facebook page by clicking on this sharing button (if you are not already connected on Facebook and you want to share the content you can either connect yourself on a separate page or you will be redirected to Facebook login page after clicking on the button).

D Complete questionnaire (translated from French)

[Part 1. Introduction and pretreatment questions]

We carry out a survey on behaviour in social networks. This survey includes a set of questions about you, your use of social networks and your political preferences. You will then be exposed to information that you could usually find in your Facebook newsfeed. The questionnaire will take about 15 minutes to be completed.

Your answers are anonymous; we will publish only aggregate results based on the survey.

Q0 Do you want to proceed with the survey?

- Yes
- No

Q1 Among the following social networks, which ones do you use?

- Facebook
- Instagram
- Twitter
- Snapchat
- LinkedIn
- None of the above

Q2 How old are you?

- Age (in years):

Q3 What is the size of the city or town where you currently live?

- 1) Less than 2500 people
- 2) Between 2500 and 5000 people
- 3) More than 5000 people
- 4) I prefer not to answer this question

Q4 What is the highest academic degree that you have?

- No diploma
- Primary school certificate
- B.E.P.C. certificate
- Certificate of professional skill (CAP)
- Certificate of vocational education (BEP)
- Vocational high school degree (BAC)
- High school degree (BAC)
- BAC + 2 (BUT, BTS, schoolteacher, DEUG, social or paramedical diploma)
- Higher education

Q5 What is your gender?

- 1) Male
- 2) Female

Q6 Your place of birth:

- 1) France
- 2) Abroad
- 3) I prefer not to answer this question

Q7 Place of birth of your father:

- 1) France
- 2) Abroad
- 3) I prefer not to answer this question

Q8 Place of birth of your mother:

- 1) France
- 2) Abroad
- 3) I prefer not to answer this question

Q9 Region of current residence:

– Region:

Q10 What is your marital status?

- 1) Single
- 2) Married
- 3) Civil union without being married
- 4) PACS (civil solidarity pact)
- 5) Divorced
- 6) Widow
- 7) I prefer not to answer this question

Q11 If you add all income sources in your household, what is your household's monthly income?

- 1) Less than 1000 euros per month
- 2) Between 1001 and 1500 euros per month
- 3) Between 1501 and 1750 euros per month
- 4) Between 1751 and 2000 euros per month
- 5) Between 2001 and 2500 euros per month
- 6) Between 2501 and 3000 euros per month
- 7) Between 3001 and 4000 euros per month
- 8) Between 4001 and 5000 euros per month
- 9) Between 5001 and 7000 euros per month
- 10) More than 7001 euros per month
- 11) I prefer not to answer this question

Q12 Do you have children?

- 1) Yes
- 2) No
- 3) I prefer not to answer this question

Q13 How often do you use Facebook?

- 1) Several times a day
- 2) Once a day
- 3) Once a week
- 4) Once a month or less
- 5) I prefer not to answer this question

Q14 How many Facebook friends do you have?

- Number of friends:

Q15 How frequently do you share content on Facebook?

- 1) Never
- 2) Seldom
- 3) Often
- 4) I prefer not to answer this question

Q16 (on a scale from 1 to 5) For you, how important are the following reasons to share content on Facebook?

- **Q16 _ 1:** I think that the content could be of interest to my friends
- **Q16 _ 2:** I want to influence my friends
- **Q16 _ 3:** I want my friends to have a good image of me
- **Q16 _ 4:** My friends make the effort of sharing content with me, I want to do the same

Q17 In order to be informed on politics, which media do you use most often?

- 1) TV
- 2) Radio
- 3) Internet
- 4) National newspapers
- 5) Local newspapers
- 6) Free newspapers
- 7) Other:
- 8) None
- 9) I prefer not to answer this question

Q18 On a scale from “1” (never) to “5” (often), comment on the following statements

- **Q18 _ 1:** I have given money to a charitable organization
- **Q18 _ 2:** I have given money to a foreigner in need (or a foreigner who asked me for money)
- **Q18 _ 3:** I have worked for a charitable organization
- **Q18 _ 4:** I have donated blood

Q19 On a scale from 1 to 5, where 1 means “it does not fit me” and 5 means “it fits me perfectly”, comment on the following statements

- **Q19 _ 1:** If someone helps me, I am ready to help back
- **Q19 _ 2:** I go out of my comfort zone in order to help someone who helped me in the past
- **Q19 _ 3:** I am ready to bear a personal cost in order to help someone who helped me in the past

Q20 On a scale from 1 to 5, where 1 means "it does not fit me" and 5 means "it fits me perfectly", comment on the following statement

- It is important for me not to be considered selfish

Q21 Can you tell me what your religion is, if any?

- 1) Catholic
- 2) Protestant
- 3) Jewish
- 4) Muslim
- 5) Buddhist
- 6) Other religion
- 7) No religion
- 8) I prefer not to answer this question

Q22 Do you usually attend religious services:

- 1) Several times a week
- 2) Once a week
- 3) Once or twice a month
- 4) Sometimes, during important festivities
- 5) Only for ceremonies, marriages, ...
- 6) Never
- 7) I prefer not to answer this question

Q23 Which are the levels of government that you feel closer to?

- 1) Europe
- 2) France
- 3) Region
- 4) Department
- 5) City/town
- 6) Local community
- 7) I prefer not to answer this question

Q24 Are you registered to vote?

- 1) Yes
- 2) No
- 3) Soon
- 4) I prefer not to answer this question

Q25 Who did you vote for in the first round of the 2017 presidential elections?

- Francois Fillion
- Marine Le Pen
- Emmanuel Macron
- Jean-Luc Melanchon
- Other candidate
- I did not vote

Q26 Who did you vote for in the second round of the 2017 presidential elections?

- Marine Le Pen
- Emmanuel Macron
- I did not vote

Q27 In general what is your opinion of the European Union?

- 1) Very positive
- 2) Positive
- 3) Neutral
- 4) Negative
- 5) Very negative

[End of Part 1.]

[Part 2. Treatments.]

See Section **C** of this Appendix.

Alt-Facts Treatment Text and questions presented in Subsection **C.3**

Imposed Fact-Check Treatment Text and questions presented in Subsection **C.4**

Imposed Fact-Check Treatment Text and questions presented in Subsection **C.5**

[End of Part 2.]

[Part 3. Post-treatment Questions]

Q32 We usually rank French people on a scale from 1 to 10 going from the left to the right of the political spectrum. How would rank yourself on this scale?

- Political orientation (ten-point scale):

Q33 Are you likely to vote for any of the following party lists? ("1" not very likely, "5" very likely)

- **Q33_1:** La France Insoumise
- **Q33_2:** LREM (Renaissance supported by La Republique En Marche, Le Modem and their allies)
- **Q33_3:** RN (Prenez le Pouvoir, list supported by Marine Le Pen)
- **Q33_4:** UMP (union of the right and the center)
- **Q33_5:** Europe Ecologie
- **Q33_6:** A list organised by the "Yellow Vests" movement

Q34 On a scale from “1” (strongly against) to “5” (strongly in favour), are you in favour of France leaving the European Union?

– Support for France leaving the EU (1 to 5 answer):

Q35 Do you think that France gains or loses from its membership in the EU? (on a scale from “1” loses to “5” gains)

– France gains from its EU membership (1 to 5 answer)

Q36 From your perspective, does the European project constitute a source of hope or a source of concern?

- 1) Neither of them
- 2) A source of hope
- 3) A source of concern
- 4) I prefer not to answer this question

Q37 Are you concerned by the fact that France does not control its borders and laws anymore due to its membership in the European Union?

- 1) Not at all concerned
- 2) A little concerned
- 3) Very concerned
- 4) I prefer not to answer this question

Q38 In your opinion, who has the most to lose from Brexit, i.e., the UK’s leaving the European Union?

- 1) Everyone
- 2) The UK
- 3) The other countries in the EU
- 4) Nobody
- 5) I do not have an opinion on the matter
- 6) I prefer not to answer this question

Q39 In your opinion, which are the two major problems that the EU is currently facing?

- 1) Purchasing power
- 2) Immigration
- 3) Security and fight against terrorism
- 4) Climate change
- 5) French sovereignty and French identity
- 6) Health
- 7) Taxation
- 8) Unemployment
- 9) Education

Q40 Among the following issues, which are the ones that will be most relevant for your vote at the European elections? (choose up to 2 options)

- 1) Purchasing power
- 2) Immigration
- 3) Security and fight against terrorism
- 4) Climate change
- 5) French sovereignty and French identity
- 6) Health
- 7) Taxation
- 8) Unemployment
- 9) Education

Q41 In your opinion, what is the percentage of French laws coming from European directives?

- 1) 0%
- 2) 10%
- 3) 20%
- 4) 30%
- 5) 40%
- 6) 50%
- 7) 60%
- 8) 70%
- 9) 80%
- 10) 90%
- 11) 100%

Q42 We are going to present you a list of sources. For every source you can specify the degree of confidence that it inspires you: a lot, enough, not much, not at all

- **Q42_1**: National newspapers
- **Q42_2**: Local newspapers
- **Q42_3**: Online newspapers
- **Q42_4**: Fact-checking websites
- **Q42_5**: TV
- **Q42_6**: OECD
- **Q42_7**: Eurostat

[End of Part 3.]

E Definitions of variables

E.1 Measuring sharing

For each URL, Google Analytics provided us with hourly data such as the number of visits, the location (city), and the share of visitors who did not interact with the page i.e., the “bounce rate”.²⁶ We merged the survey data and the data from Google Analytics by hour of the day and treatment.

This matching allowed us first of all to measure the probability that a participant visited the survey. The variable *visit the webpage* takes the value

The other goal was to measure the probability that someone actually shared on Facebook and the total number of reshares by friends of participants. To measure the number of shares on Facebook by the participants we use the bounce rate indicated by Google Analytics and the number of entrances in the website:

$$\text{Number of shares}_{Survey} = (100 - \text{Bounce Rate}) \times \text{Entrances}$$

Note that we use the number of entrances instead of visits because the bounce rate is computed within a session such that someone has to enter the website via a specific page and if he did not interact with the page the session ends on the same page and the bounce rate is 100%. The number of entrances can differ from the number of unique visits since people could also view the fact check or open both the article from the survey and the one shared on Facebook which will count as one entrance on the landing page only. In addition, since the data are aggregated per hour, if someone stays on the same page for more than one hour and refreshes it, it will count as one unique view in both hours but only one entrance.

Finally, the number of reshares can be calculated using the total number of shares from Facebook via the share buttons with the count option. From this total number of shares, we just subtract the shares by the participants, as calculated above:

$$\text{Number of shares}_{Facebook} = \text{Number of shares}_{Total} - \text{Number of shares}_{Survey}$$

Since the only interactive component of the webpage was the share button, we can use the bounce rate to infer the number of visitors who clicked on the share button.

²⁶The bounce rate is equal to 100% for a visitor who came to the page and closed it without clicking anywhere on the page during his session.

E.2 Construction of variables measuring behavioral traits and motives to share

Social desirability score:

Social desirability score is defined as the first principal component of answers to the following three questions in the survey: “I have worked for a charitable organization,” “I have donated blood,” and “It is important for me not to be considered selfish.” For each of these questions the participants gave an answer on a scale from 1 to 5 (“never” to “often” for the first two questions, and “it does not fit me” to “it fits me perfectly” for the third question). The factor loadings are: 0.6493 (worked for charity); 0.6412 (blood donor); 0.4090 (importance of not being considered selfish).

Other behavioral traits:

Self-reported altruism was measured as the average response to the following four questions (based on the self-report altruism scale first proposed by [Rushton et al. \(1981\)](#)) divided by 5 (since the responses are on a 1-5 Likert scale)

- I have already given money to a charitable organization
- I have already given money to a foreigner in need (or that asked me for money)
- I have already worked for a charitable organization
- I have already donated blood

Donor of blood and money was measured as the average response to the following four questions (based on the self-report altruism scale first proposed by [Rushton et al. \(1981\)](#)) divided by 5 (since the responses are on a 1-5 Likert scale)

- I have already given money to a charitable organization
- I have already given money to a foreigner in need (or that asked me for money)
- I have already donated blood

Self-reported reciprocity was measured as the average response to the following three questions (adapted from [Dohmen et al. \(2008\)](#)) divided by 5 (since the responses are on 1-5 Likert scale)

- If someone helps me, I am ready to help her back
- I go out of my comfort zone in order to help someone who helped me in the past
- I am ready to bear a personal cost in order to help someone who helped me in the past

Importance of self-image was measured as the response to the following question (adapted from [Henry and Sonntag \(2019\)](#)) divided by 5 (since the responses are on 1-5 Likert scale)

- It is important for me not to be considered selfish

Motivations for sharing: We asked the following question before the treatment: “For you, how important are the following reasons to share content on Facebook?” with answers on a 1-5 scale. The answer to each option (divided by five) defines a motive.

- I think that the content could be of interest for my friends (**Interest**)
- I want to influence my friends (**Influence**)
- I want my friends to have a good image of me (**Image**)
- My friends make the effort of sharing content with me, I want to do the same (**Reciprocity**)

F Theoretical model

F.1 Analytical framework: checking and sharing

Our empirical results in Section II.D show that Viewers have a much higher ex ante propensity to share alt-facts than Nonviewers. This may seem puzzling: those who would share alt-facts regardless of the information contained in the fact-check should have no incentive to view, thus implying a higher propensity to share for Nonviewers relative to Viewers.

In this section, we develop a theoretical framework to shed light on these results. We jointly model strategic choices of viewing fact-checking and of subsequent sharing of alt-facts. (For the sake of simplicity, we start with a version of the model where we assume away the option of sharing fact-checking.) The key idea is that viewing is costly, so for individuals who have either very high or very low propensity to share, regardless of the evidence, there is no incentive to incur the cost of viewing to collect the evidence. Thus only those with intermediate propensity to share alt-facts are likely to view the fact-checking. The model formalizes this idea and derives conditions under which among those who choose not to view, those who tend not to share alt-facts are more numerous, either because of the underlying heterogeneity in types or because of the way the fact-checking is designed.

We consider a group of individuals i who receive a piece of news, that could be true or false. The state of the world $s \in \{0, 1\}$ corresponds to the veracity of the news, where $s = 1$ indicates that the news is true. To simplify the exposition, we assume that all members of the group initially have the same prior beliefs: they expect that the state is $s = 1$ with probability q and $s = 0$ with probability $1 - q$.

Individuals choose whether to share the news. They are heterogeneous in terms of returns to sharing. The payoff of sharing V_i^s is state-dependent. The net value of sharing is positive when the news is true and negative when it is false $V_i^0 < 0 < V_i^1$. The values of V_i^s could reflect different motives for sharing identified in the empirical results above, such as the intention to influence (see Section F.2 for the formalization of these different motives).

Before sharing each individual has the opportunity to view fact-checking at cost c . Viewing is expected to produce an ex post belief which is either (i) $q^- < q$ (with probability p), i.e a report suggesting that the state is probably low, or (ii) $q^+ > q$ (with probability $1 - p$). Bayes plausibility implies

$$q = pq^- + (1 - p)q^+. \quad (1)$$

To simplify the exposition, we assume here that there is no heterogeneity in their level of trust in the fact-checking.²⁷

²⁷Differential level of trust in fact-checking sources can be modeled as differences in $q^+ - q^-$. In particular an individual who has no trust in fact checking sources would have $q^+ = q^- = q$ and would never choose to view fact-checking — as viewing is costly and she expect no additional information from viewing.

Sharing. Individual i with belief q shares alt-facts if and only if

$$qV_i^1 + (1 - q)V_i^0 \geq 0 \Leftrightarrow q \geq q_i^* \equiv \frac{-V_i^0}{\Delta V_i}.$$

where $\Delta V_i \equiv (V_i^1 - V_i^0)$.

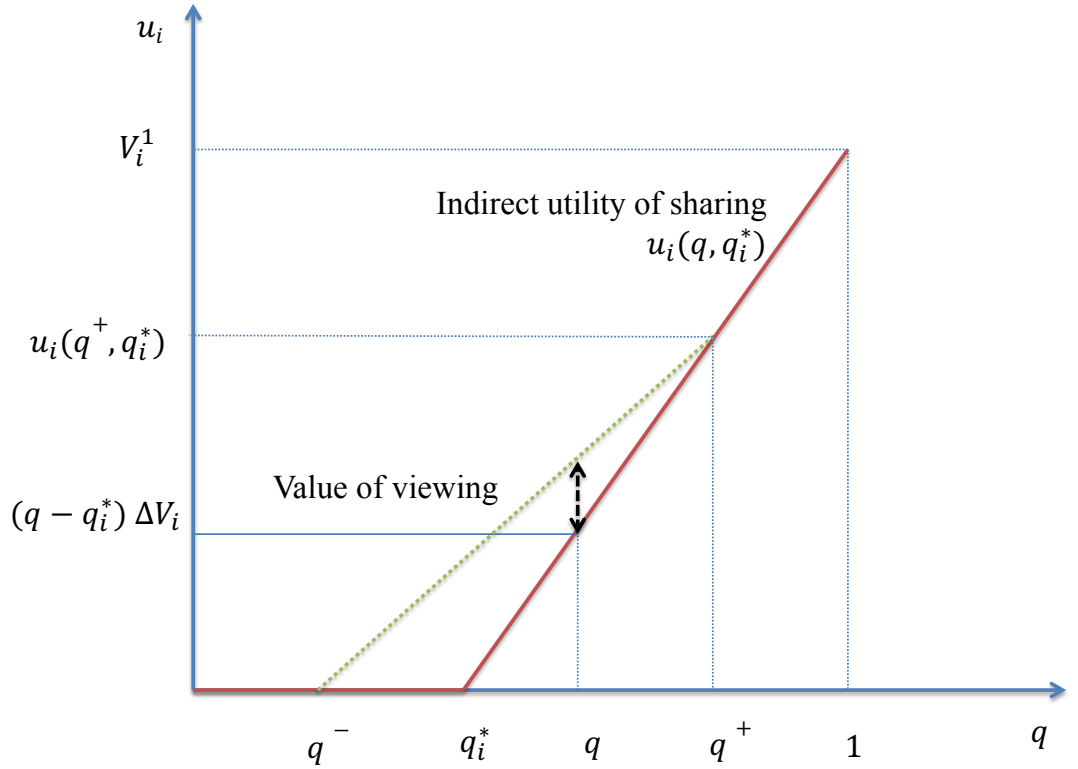
Individual i shares if and only if q is higher than an individual-specific threshold q_i^* . The threshold q_i^* therefore characterizes the individual's propensity to share. We assume q_i^* is distributed on $[0, 1]$ with a cumulative distribution function $F(\cdot)$.

The indirect utility of individual i is thus given by

$$u(q, q_i^*) = \max\{0, qV_i^1 + (1 - q)V_i^0\} = \max\{0, (q - q_i^*)\Delta V_i\}. \quad (2)$$

We plot the value of sharing in Figure F1 (solid line). Individual i does not share for $q \leq q_i^*$ and gets zero value while the expected value of sharing is then linearly increasing in q for $q > q_i^*$.

Figure F1: Decision to view fact-checking



Note: The solid line represents the indirect utility of sharing alt-facts $u_i(q_i, q_i^*)$. The dotted line is a straight line connecting points $q^-, u_i(q^-, q_i^*)$ and $q^+, u_i(q^+, q_i^*)$. The distance between two lines represents the value of viewing fact-checking (gross of the cost of viewing c) for a given q .

Viewing. We now turn to the decision of whether to view the fact-checking information prior to sharing. The utility of individual i who does not view is given by expression (2). If the individual does view fact-checking, she expects to get, with probability p , a report q^- suggesting that the state is probably low and with complementary probability $1 - p$ a report $q^+ > q$ suggesting it is likely to be high. The cost of viewing is c . The expected utility of a viewer is therefore given by:

$$-c + pu(q^-, q_i^*) + (1 - p)u(q^+, q_i^*).$$

This immediately implies that if individual i has extreme beliefs ($q_i^* < q^-$ or $q_i^* > q^+$), she does not view the fact-checking. Indeed, if $q_i^* < q^- < q < q^+$ she knows that either with or without viewing she would always share; therefore it makes no sense to pay a non-trivial cost of viewing. Similarly, if $q_i^* > q^+ > q > q^-$, then she knows that she would never share — so no need to view.

For those with intermediate propensity to share $q_i^* \in (q^-, q^+)$, the incentives to view the fact-checking depend on the cost of viewing c as characterized in Proposition 1.²⁸

Proposition 1. *Individual i views the fact-checking if and only if $q_i^* \in (\underline{q}_i, \bar{q}_i)$, where $\underline{q}_i = \min \left\{ q, q^- + \frac{c}{p\Delta V_i} \right\}$ and $\bar{q}_i = \max \left\{ q, q^+ - \frac{c}{(1-p)\Delta V_i} \right\}$.*

For the sake of simplicity let us now consider the case where $\Delta V_i = \Delta V$ for all i . In this case, there is no variation in the slope of the value function ΔV_i , so that variations in q_i^* are driven by variations in the intercept V_i^0 ; the range $(\underline{q}_i, \bar{q}_i)$ is the same for all i . We maintain this assumption below and denote this range (\underline{q}, \bar{q}) .

Proposition 1 allows us to compare the average ex ante propensity to share of those who decide to view (Viewers) and those who do not (Nonviewers). The probability that a Nonviewer shares is given by $\frac{F(q)}{F(q) + (1 - F(\bar{q}))}$, while for a Viewer the probability of sharing is $\frac{F(q) - F(\underline{q})}{F(\bar{q}) - F(\underline{q})}$.

Nonviewers are made up of two distinct groups: those who do not view because they never share or almost never share (very low ex ante propensity), and those that do not view because they share always or almost always (very high ex ante propensity). Whether the average sharing among Nonviewers is higher than among Viewers depends on the relative size of the two groups.

²⁸ In Figure F1 the value of viewing (gross of viewing cost c) $pu(q^-, q_i^*) + (1 - p)u(q^+, q_i^*)$ is represented by the dotted line, which is the average between the zero payoff of not sharing if the report is q^- and the positive payoff if the report is q^+ . The intuition is that the indirect utility of sharing (the solid line in Figure F1) is weakly convex. So the individual receive a strictly positive gain from viewing whenever $q^- < q_i^* < q^+$. Indeed, in this case, viewing represents a lottery (q^- with probability p and q^+ with probability $1 - p$) while not viewing is a certain outcome $q = pq^- + (1 - p)q^+$ (see (1)). If $q_i^* < q^-$ or $q_i^* > q^+$ then the individual faces a linear part of the indirect utility function. In this case, gain from viewing is precisely zero; as there is a non-trivial cost $c > 0$, she never views.

Proposition 2. *Viewers have a higher ex ante propensity to share than Nonviewers if and only if*

$$\frac{F(\underline{q})}{1 - F(\bar{q})} < \frac{F(q)}{1 - F(q)}. \quad (3)$$

There are two key parameters in (3): the distribution $F(\cdot)$ of propensity to share and the design of the fact-checking technology, that determines q^- and q^+ , and thus ultimately \underline{q} and \bar{q} . Condition (3) is more likely to be satisfied if the fact-checking technology tends to disprove statements, i.e., q^- or/and q^+ low, if the distribution of propensity to share has a large mass at 1 (i.e. $1 - F(\bar{q}) \simeq 1 - F(q)$) or if the prior q is high.

Let us first consider the distribution function $F(q)$. A large mass at $q = 1$ implies that there are many individuals who would never share (53% of our sample report never sharing or sharing seldom). The fact-checking technology is also important. For example, condition (3) is more likely to be satisfied when both ex post signals q^- and q^+ are low (which is the case when there is a high chance of getting message q^+ , see (1)). This is the case where fact-checking either produces a conclusive signal that the information is false (q^- is close to zero) or an inconclusive signal stating that no evidence was found to disprove it (q^+ is substantially below 1).²⁹ The individuals who benefit from such a technology are those who are likely to share and would only choose not to share if they received a signal that showed without doubt that the state was low. Thus, when fact-checking produces such signals, high propensity sharers are induced to view.

Sharing fact-checking.

In this subsection, we extend the model adding an opportunity for the Viewers to share fact-checking as well. We denote the payoff of sharing fact-checking $V_i^{s,FC}$. The value of sharing fact-checking when the news is true is smaller than when it is false $V_i^{1,FC} < 0 < V_i^{0,FC}$. We also denote $\Delta V_i^{FC} = V_i^{0,FC} - V_i^{1,FC} \geq 0$.

This is the mirror case of sharing alt-facts. Individual i shares the fact-checking if and only if the belief that the news is true is low enough. Specifically, individual i shares if and only if q is lower than an individual specific threshold $q_{i,FC}^*$. We assume that $q_{i,FC}^*$ is drawn from a distribution $G(\cdot)$ with support on $[0, 1]$. We further assume that the draw of q_i^* , that determines the sharing of alt-facts, and the draw of $q_{i,FC}^*$, that determines the sharing of fact-checking are independent. The model can easily be extended to the case where the draws are correlated (capturing for instance the idea that some individuals have a general taste for sharing), at the cost of notational complexity.

The (additional) indirect utility from sharing fact-checking is given by

$$u_i^{FC}(q, q_{i,FC}^*) = \max\{0, (q_{i,FC}^* - q)\Delta V_i^{FC}\}. \quad (4)$$

²⁹This is a good approximation of how fact-checking works in reality: fact checkers either succeed (and produce report $q^- = 0$) or fail to disprove the original statement (and produce an inconclusive report).

However, opportunity to share fact-checking influences the viewing decision in a different way. The reason is that without viewing there is no possibility to share the fact-checking (whereas sharing alt-facts was always a possibility). Viewing thus provides an additional expected benefit due to the possibility to share the fact-checking. This benefit is given by:

$$B^{FC}(q_{i,FC}^*) = pu^{FC}(q^-, q_{i,FC}^*) + (1-p)u^{FC}(q^+, q_{i,FC}^*) \geq 0 \quad (5)$$

This additional benefit is increasing in $q_i^{*,FC}$, so that on average, those who view have a higher ex ante propensity to share fact-checking. For the viewing decision this acts as a shifter of costs, so that Proposition 2 naturally extends to the following result:

Proposition 3. *Viewers have a higher ex ante propensity to share fact-checking. Furthermore, individual i views the fact-checking if and only if $q_i^* \in (\underline{q}_i, \bar{q}_i)$, where $\underline{q}_i = \min \left\{ q, q^- + \frac{c-B^{FC}(q_{i,FC}^*)}{p\Delta V_i} \right\}$ and $\bar{q}_i = \max \left\{ q, q^+ - \frac{c-B^{FC}(q_{i,FC}^*)}{(1-p)\Delta V_i} \right\}$.*

Overall, the model shows that viewers have a higher propensity to share fact-checking. They also have a higher propensity to share alt-facts under conditions specified in Proposition 2, conditions that appear reasonable in our empirical setting.

F.2 Formalization of sharing motives

Here we present different interpretations of the value of sharing V_i^0 and V_i^1 :

- If individual i shares to influence, she knows that if the state is $s = 0$, the news is less likely to influence the recipients, who could for instance receive fact-checking from other sources. It is thus natural to have $V_i^1 > V_i^0$. Sharing false news might even backfire, implying $V_i^0 < 0$.
- If individual i has image concerns, she might care about inferences others make when she shares news that turns out to be false, inferences about either her motives or her judgment. Denote $-R$ the negative reputation obtained if i has shared fake news. Then $V_i^0 = V_i^1 - \mu_i R$, where μ_i measures the strength of image concerns.
- If individual i faces moral costs of sharing content she believe has some chance of being false (something we discuss more in depth in the next section), the belief q will matter. Specifically we assume that sharing provides a payoff V and has an individual specific cost $c_m^i(1-q)$ where c_m^i is the moral cost of sharing and $(1-q)$ is the belief that the state is wrong. This is exactly the model above with $V_0^i = V - c_m^i$ and $V_1^i = V$.

F.3 Proofs

Proof of Proposition 1

If $q_i^* < q^-$, even if the individual receives the low report q^- , she would still share, so there is no value of checking. Similarly for $q_i^* > q^+$, i never shares, so there is no value in checking. This proves the first part of the proposition.

We now turn to the second part.

Suppose $q^- \leq q_i^* < q$, so that without checking, the individual shares and gets value

$$u_{nc} = (q - q_i^*)\Delta V_i$$

If she checks, she does not share when she receives q^- (with probability p) and shares when receiving q^+ (happens with probability $1 - p$).

$$u_c = (1 - p)(q^+ - q_i^*)\Delta V_i - c$$

Thus the net value of checking is

$$\Delta V_i [(1 - p)q^+ - q + pq_i^*] - c$$

If ΔV_i is constant, i checks if

$$q_i^* \geq \frac{q - (1 - p)q^+}{p} + \frac{c}{p\Delta V} = q^- + \frac{c}{p\Delta V}$$

Similarly, if $q_i^* > q$, the net value of checking is

$$\Delta V_i [(1 - p)(q^+ - q_i^*)] - c$$

If ΔV_i is constant, i checks if

$$q_i^* \leq q^+ - \frac{c}{(1 - p)\Delta V}$$

Proof of Proposition 2

A direct consequence of Proposition 1 is that the probability that a non-checker shares is given by $\frac{F(\underline{q})}{F(\underline{q})+(1-F(\bar{q}))}$, while for a checker it is given by $\frac{F(\underline{q})-F(\bar{q})}{F(\bar{q})-F(\underline{q})}$. The condition can be expressed as:

$$\frac{F(\underline{q})}{1-F(\bar{q})} < \frac{F(\underline{q})}{1-F(\underline{q})} \quad (6)$$

The left hand side is increasing in \underline{q} and \bar{q} while the right hand side is increasing in \underline{q} , implying the comparative statics in Proposition 2.