Economics Bulletin

Volume 42, Issue 2

Fake news and beliefs: evidence from a natural experiment

Francesco Bogliacino
Universidad Nacional de Colombia

Rafael Charris
Chapman University

Camilo Gómez
Universidad Nacional de Colombia; CERGE-EI

Felipe Montealegre
Università degli Studi di Bologna

Abstract

Can fake news affect individual beliefs? We exploit the natural experiment of 'the night of collective panic' in Bogotá, combining administrative data from emergency calls, tweets, and elicited beliefs, to answer this question. The outcome variables are the perceived levels of insecurity within the neighbourhood. We fail to reject the hypothesis that the perception of insecurity was unaffected. We cannot exclude that the sample's demographic characteristics drive the null result.

Bogliacino: Universidad Nacional de Colombia and Centro de Investigaciones para el Desarrollo (Corresponding author, Facultad de Ciencias Económicas, Universidad Nacional de Colombia, Kr 30, No 45-03, Bogotá (Colombia), fbogliacino@unal.edu.co). Charris: Chapman University, rcharris@chapman.edu, Gómez: Centro de Investigaciones para el Desarrollo and CERGE-EI, caegomezca@unal.edu.co. Montealegre: University of Bologna, andres.montealegre@studio.unibo.it. Funding from Universidad Nacional (Convocatoria Jesus Antonio Bejarano 2017). The authors appreciate comments and suggestions from participants at various seminars. Comments from the Editor and two anonymous referees have been very useful.
Citation: Francesco Bogliacino and Rafael Charris and Camilo Gómez and Felipe Montealegre, (2022) "Fake news and beliefs: evidence from a natural experiment", *Economics Bulletin*, Volume 42, Issue 2, pages 802-809
Contact: Francesco Bogliacino - fbogliacino@unal.edu.co, Rafael Charris - rcharris@chapman.edu, Camilo Gómez - caegomezca@unal.edu.co, Felipe Montealegre - andres.montealegre@studio.unibo.it.
Submitted: April 02, 2022. **Published:** June 30, 2022.

1. Introduction

The spread of blatantly false or inaccurate information has come under increasing scrutiny. The diffusion of new media platforms in recent years has lowered the entry cost for the provision of information (Zhuravskaya et al., 2020), jeopardizing the binding nature of the self-regulatory norms of traditional media (Lazer et al., 2018) and thrusting the level of polarization via the 'echo chamber' effect.[1]

How does fake news work? False news may usher a change in beliefs. Over-exposure to an event that captured the media's attention affects judgment, as people can fall victim to the availability heuristics (Tversky and Kahneman, 1974). The availability heuristics is the mental shortcut that leads to evaluations based on the ease of recalling information. In this context, fact-checking may be ineffective (Barrera et al., 2020) because fake news campaigns involve 'social identity' (Tajfel, 1978). Empirical information that contradicts identity beliefs may be involuntarily discarded ('motivated reasoning'; Kahan et al., 2013), as doubting those beliefs threatens the attachment to a group.

Although this fake news-belief causal nexus is firmly posited (Bago et al. 2020), there is no established consensus (Della Vigna and Gentzkow, 2010; Pennycook et al. 2018). The present work contributes to this debate by using a novel source of quasi-experimental variation.

The context is as follows. During a national strike on the evening of the 21st November 2019 in Colombia, social media platforms were spreading videos of (supposed) violations of private property and individuals shooting guns into the air to frighten alleged looters. The day after, Bogotá was the epicentre of this campaign, and the Colombian government decreed a curfew to regain control of the situation. The Police and the Major later confirmed that the unprecedented number of emergency calls did not result in any concrete act of vandalism or any threat whatsoever.

It is plausible that this campaign worsened the perception of insecurity. Taking advantage of a few lab and artefactual field experiments conducted before the event, over the period from August 2018 to November 2019, we reached out to a sample of participants to fill in an online questionnaire in the week following the night of collective panic. Since the original protocol

[1] It refers to the fact that members match with persons with similar beliefs or opinions, in online communities, reinforcing existing views and neglecting alternative ideas.

included three questions on the perceptions of insecurity within the neighbourhood, we elicited them again to measure their change.

Our identification strategy relies on standard differences-in-differences. Using data from emergency calls, participants are split into a treatment and a control group according to the level of exposure. Complementary information extracted from tweets is used to validate the level of exposure to fake news.

2. Materials and methods

2.1 Beliefs elicitation

An invitation was sent to 774 persons, to participate in an online experiment. The sample was drawn from the participants to three previous experiments, conducted in August 2018, September-November 2018, and October-November 2019, complemented with a control group, who faced those questions for the first time. Data collection was limited to seven days to avoid a spillover. A total of 448 subjects answered (58% response rate). We excluded ten subjects because of a coding error. The final sample consisted of 438 participants, with 40% female and the average age around 22 years old (+/- 3). The placebo sample includes 82 subjects (18.72%).

The experiment was programmed with oTree (Chen et al., 2016), and the complete experimental protocol is available in SOM, Section I.

Since the pre-shock data were collected in a post-experimental questionnaire after an incentivized task, we used a (distractor) task with incentivized questions, to avoid experimenter demand. After this task, participants filled in a short questionnaire, where we elicited our main outcome variables.

The three questions read as follows: (1) "The neighbourhood where I live is violent" [1-5]; (2) "Do you feel safe walking alone in your neighbourhood at night? [Yes=1]"; (3) "In case of need, would you call for help from an unknown person in your neighbourhood? [Yes=1]".

For question (1), we have 195 matched answered pre- and post-shock. For questions (2) and (3), we have 320 matched answered pre- and post-shock. For 159 subjects, we have matched answered for the three questions.

We collected data on participants' gender, age, Socio Economic Strata number (Bogliacino, Jiménez Lozano, & Reyes, 2018), district of residence, level of trust, and emotional reactions during the general strike of the 21st of November.

2.2. Data on exposure to false and fake news

In Colombia, there is in force an information act, which permits access to public information except if it involves confidentiality or national security. We sent a request to public authorities to access the data on emergency calls during the days of the 21st and 22nd November 2019 in Bogotá, by district, related to looting, robbery, or property invasion (the list of categories is in SOM Section III). The treated group consists of those living in the nine districts where the number of calls is strictly above the median. We assign the rest to the control group. As confirmed by the police.[2] These calls were not associated with any real threat, thus differences in exposure cannot be ascribed to differences in crime rate.

2.3 Tweet data

We also collected tweets to validate exposure to fake news. The keywords used to find the tweets were the names of the districts in Bogotá, plus the first 16 trending topics in the city for 21st and 22nd November 2019.

3. Results

The 'panic attack' was indeed a significant episode. In Figure 1, we report the Google® Trends data for the search "Vándalos" (vandals). Notice that the search peaked on the date of the panic attack at a level 10-20 times the average activity during the previous year.

Additionally, on Twitter®, six out of 10 trending topics were directly or indirectly related, and the event covers 5.23% of the total number of tweets in Colombia over the 48 hours.[3]

[2] Cifras y Conceptos (2020) used the Information Act to acquire the relevant data from the Police. There was (at the time of the request) no active police investigation over looting in apartments that night, nor any judicial measures against anybody.

[3] A wordcloud is reported in the SOM Figure A.1, where we also report a table with the overall incidence of the words "vandalo, metiendo, conjunto", which refer to the expression "Someone is entering the adjacent building" that was popularized during the night of panic (SOM, Table A1).

Figure 1: Google trends in Colombia for "Vandalos" (vandals), max=100.

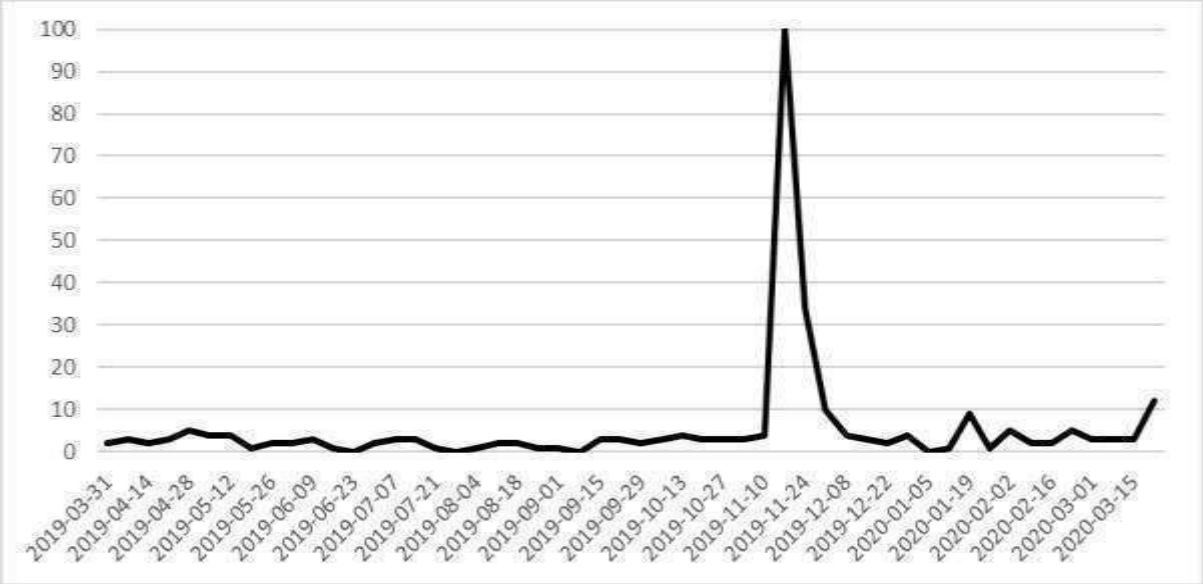
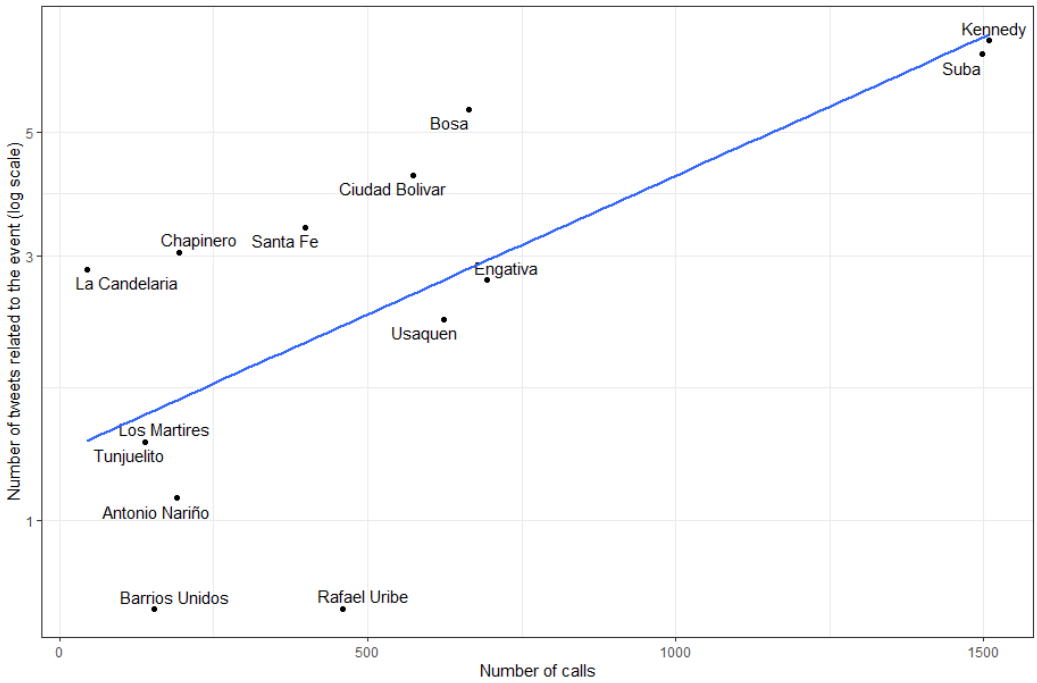


Figure 2 below plots the data on tweets and emergency calls by district to validate the measure of exposure to fake news (Spearman's ρ =0.62, p=0.02).

Figure 2: Number of tweets related to the event and number of calls by district.



Third, we assess the validity of the responses, by comparing *ex-post* answers by matched participants and by placebo group. We fail to reject the null hypothesis (for question (1), χ^2=4.56, p=0.33; (2), χ^2=0.99, p=0.31; (3), χ^2=0.00, p=0.99).

The critical identification assumption is that of parallel trends.[4]

When we perform the differences-in-differences analysis, we do not detect any effect of the shock. The relevant coefficient is the interaction between the ex-post and the treatment dummy. The interpretation of the coefficient is the difference in the variation of the beliefs for the treated group with respect to control subjects. In Table I, we report the results. For question (1), the statistics is t=-0.19, p=0.848, for question (2) the statistics is t=0.18, p=0.853, and finally for question 3, the statistics is t=-0.46, p=0.644. We also report the results restricted to the sample for which we have all the questions, but the results do not change.

We performed various robustness checks: a) adding district-level fixed effects; b) controlling for the size of the district; c) controlling for the level of trust, the participation in the strike and the emotional reaction during the event; d) replacing the 60th percentile to the median; e) running the dif-in-dif with continuous treatment (number of calls and number of related tweets) following Acemoglu et al. (2004). (The latter estimations should eliminate the concern that this is due to low statistical power.) The results are virtually unaffected (SOM, Table A2-A8 and Figure A.3).

Table I. Mean difference with respect to control on standardized and per-task accuracy, response time and efficiency conditional on several exogenous control variables.

	(1)	(2)	(3)	(4)	(5)	(6)
	Do you agree with the statement "The neighbourhood where I	Do you feel safe walking alone in your neighbourhood at night? [Yes=1]	In case of need, would you reach for help to an unknown person in	Do you agree with the statement "The neighbourhood where I	Do you feel safe walking alone in your neighbourhood at night? [Yes=1]	In case of need, would you reach for help to an unknown person in

[4] As documented in the SOM, it seems plausible. Since we have different points in time, we can look at pre-trends. If we perform Mann Whitney Wilcoxon rank sum tests on the *ex-ante* data period of collection, we found that for question (1), the *ex-ante* trend is non-existent for both treated individuals (z=0.74, p=0.45) and for the control group (z=-1.45, p=0.14). This is equally the case for question (2), where there is lack of trend for both groups: for treated individuals z=0.44, p=0.65 and for the control group z=1.51, p=0.12. For question (3), for treated individuals, there is a time variation before the shock (z=2.27, p=0.02), but not for the control group (z=1.04, p=0.29), however they point to the same direction (declining trend) and there may be an issue of power for the latter group. We plot the pre-trends in SOM Figure A.2.

	live is violent" [1-5]		your neighbourhood? [Yes=1]	live is violent" [1-5]		your neighbourhood? [Yes=1]
High exposure	-0.12	0.03	-0.02	-0.17	0.08	-0.06
	(0.18)	(0.05)	(0.05)	(0.20)	(0.08)	(0.08)
Ex post	-0.30***	0.06	0.01	-0.36***	0.08	0.09
	(0.11)	(0.04)	(0.05)	(0.11)	(0.05)	(0.06)
High exposure X *ex post*	-0.03	0.01	-0.03	0.06	-0.06	-0.05
	(0.14)	(0.05)	(0.06)	(0.14)	(0.06)	(0.08)
Constant	3.13***	0.26	0.70***	3.34***	0.54	1.05***
	(0.98)	(0.18)	(0.18)	(1.04)	(0.38)	(0.32)
Obs	390	640	640	318	318	318
R-squared	0.13	0.08	0.02	0.15	0.07	0.06
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Common sample				Yes	Yes	Yes
N	195	320	320	159	159	159

Note: Differences-in-differences regression. Clustered SE at the individual level reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4. Discussion and concluding remarks

Our results suggest that beliefs do not change after intense exposure to fake news.

The main limitation is related to the use of young participants. Although our sample is not statistically different from the rest of our experimental subject pool,[5] their socio-demographic characteristics have a very limited range of variation with respect to the overall population, which may raise some issues regarding the generalizability of the (null) finding. Notice that young subjects are also more likely to use social media.

Responses are not incentivized, but the perception of insecurity is a strong predictor of generalized trust (Blanco and Ruiz, 2013).

References

[5] Since the subject pool comes from a public university is much more representative of the overall population 18-24. Additionally, one of the experiments was artefactual with participants from all the districts (Bogliacino et al., 2020).

Acemoglu, D., Autor, D.H. and Lyle, D. (2004). Women, war, and wages: The effect of female labor supply on the wage structure at mid-century. Journal of Political Economy, 112(3), 497-551

Bago, B., Rand, D. G., & Pennycook, G. (2020, January 9). Fake News, Fast and Slow: Deliberation Reduces Belief in False (but Not True) News Headlines. Journal of Experimental Psychology: General.

Barrera, O., Guriev, S., Henry, E., & Zhuravskaya, E. (2020). Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics*, 182:104-123

Blanco, L., & Ruiz, I. (2013). The impact of crime and insecurity on trust in democracy and institutions. In *American Economic Review*. 103:284-288

Bogliacino, Francesco, Camilo E. Gómez, and Gianluca Grimalda. 2020. "Crime-related Exposure to Violence and Social Preferences: Experimental Evidence from Bogotá." SocArXiv. July 18. doi:10.31235/osf.io/wqst7.

Bogliacino, F., Jiménez Lozano, L., & Reyes, D. (2018). Socioeconomic stratification and stereotyping: lab-in-the-field evidence from Colombia. *International Review of Economics*, *65*(1)

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*,:88–97.

Cifras y Conceptos (2020) Persiguiendo fantasmas. Accessed on the 2021/11/25.

Della Vigna, S., & Gentzkow, M. (2010). Persuasion: Empirical Evidence. *Annual Review of Economics*, *2*(1):643–669.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2013). *Motivated Numeracy and Enlightened Self-Government*. Behavioural Public Policy, 1:54-86

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … Zittrain, J. L. (2018). The science of fake news: Addressing fake news requires a multidisciplinary effort. *Science*, *359*(6380):1094–1096

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. Journal of Experimental Psychology: General, 147(12), 1865–1880.

Tajfel, H. (1978). Interindividual Behaviour and Intergroup Behaviour. *Differentiation between Social Groups: Studies in the Social Psychology of Intergroup Relations*.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases.

Science, *185*(4157):11124–11131.

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*. 12:415-438