

7(3):90–109.

Waisbord, Silvio. 2018. “Why populism is troubling for democratic communication.” *Communication Culture & Critique* 11(1):21–34.

Waldner, David and Ellen Lust. 2018. “Unwelcome change: Coming to terms with democratic backsliding.” *Annual Review of Political Science* 21:93–113.

Whitten-Woodring, Jenifer. 2009. “Watchdog or lapdog? Media freedom, regime type, and government respect for human rights.” *International Studies Quarterly* 53(3):595–625.

Whitten-Woodring, Jenifer and Douglas A. Van Belle. 2017. “The correlates of media freedom: An introduction of the Global Media Freedom Dataset.” *Political Science Research and Methods* 5(1):179–188.

Fake News, Propaganda, and Lies Can Be Pervasive Even If They Aren’t Persuasive

by Andrew T. Little¹

University of California, Berkeley

The truth has taken a beating recently. Leading up to the 2016 election in the United States, social media was flooded with widely shared false stories. Once in office, the *Washington Post* estimated that Donald Trump made around 6 false or misleading claims per day in his first year and a half in office (Kessler, Rizzo and Kelly, 2018). Acting as a legal advisor for Trump, former New York Mayor Rudy Giuliani went so far as to proclaim that “the truth is not the truth.” Across the pond, pro-Brexit campaigners traveled around on a bus touting a misleading claim that money saved by leaving the European Union could be spent on healthcare. And the lies told by democratic politicians generally pale in comparison to their more autocratic counterparts, who some argue now use propaganda and other forms of information manipulation rather than violence as the key technology to retain power (Guriev and Treisman, 2015; Gehlbach, Sonin and Svobik, 2016).

These observations have led to considerable hand-wringing. A review of several recent books on the theme observed that “in the publishing world circa summer 2018, the death-of-truth brigade is rivaled only by the death-of-democracy crew” (Lozada, 2018).

Here I present some reasons to resist extreme pessimism. I primarily focus on the United States since the 2016 election, but draw on a combination of work about persuasion (often U.S. based) as well as recent work in comparative politics, particularly the study of information manipulation in authoritarian contexts. Two general themes from these literatures and related work are (1) manipulating beliefs is hard, and having the ability to manipulate information may backfire; and (2) actions that appear to be about manipulating beliefs often serve other purposes.²

The point is not to say, as the dog surrounded by a house on fire in a popular meme does, that “this is fine” (Plante, 2016). Rather, I aim to show that a key mechanism underlying much of this pessimistic writing — that readers of fake news and the like are actually persuaded — deserves a healthy dose of skepticism (Nyhan, 2018).

I. Causes for Concern

Here is the case for concern in brief. First, there is an increasing supply of bad information available to citizens. This is partly driven by changes to traditional media, such as the rise of the 24-hour news cycle and the increasing prevalence of opinion journalism rather than factual and investigative reporting (Rich and Kavanagh, 2018). New information and communication technologies such as social media have also made it easier to produce and consume false information. Second, the increased supply of information in general has arguably made it easier for people to find information that conforms to their prior beliefs (Sunstein, 2018).³

Combined, these trends are troubling on an individual level. They plausibly render citizens less informed about the costs and benefits of proposed policies and the performance of politicians in office. However, they might be even more problematic on a collective level. Even if citizens do have the exact same information, increased partisanship (‘motivated reasoning’) may lead to bigger disagreements about the performance of politicians *and* the quality and bias of news sources (Little, 2018). Perhaps worse, everyone receiving their own prior-reinforcing and dubious information makes

¹Many thanks to Matt Golder, Carlo Horz, Josh Kalla, Brendan Nyhan, Alex Matovski, and Daniel Stone for helpful comments and discussion.

²For an alternative perspective on this literature, which emphasizes when propaganda can be effective, see Horz (2018).

³While intuitively plausible, empirical evidence on this point is mixed (Barberá et al., 2015; Flaxman, Goel and Rao, 2016).

it harder to have a ‘common set of facts’ in order to make good policy and simply coexist with those who hold different political views (Rich and Kavanagh, 2018).

II. Why We Shouldn't Be That Concerned

The situation described above is bleak. Here are some theoretical and empirical reasons to avoid hitting the panic button immediately.

Persuasion is hard. If we are to be concerned about the effect of misinformation on people's beliefs, we need reason to think that they actually *believe* the misinformation. However, a general theme from empirical and theoretical work on persuasion — in the U.S. and more comparatively — is that persuasion is really hard. The notion that campaigns and other forms of persuasion may have ‘minimal effects’ is old and influential (Berelson, Lazarsfeld and McPhee, 1954), if not universally accepted.⁴

We can't infer from the prevalence of fake news and the like that opinions are being meaningfully changed or society further polarized. We simply don't have any direct evidence of fake news having a large impact on political beliefs or behavior.

A first reason to be skeptical of the potential for persuasion is that on the kinds of consequential beliefs that politicians and other actors care to manipulate — Which party should I vote for? Is [insert prominent policy proposal] a good idea? Is [insert prominent politician] trustworthy? — people have a lot of information.⁵ Even if a message on one of these topics is taken at face value, when the audience has observed hundreds or thousands of messages with similar precision the marginal impact of each one will be minimal.

To give something of a formalization, suppose a citizen is forming a belief about how well a politician is doing in office, represented by a continuous variable θ , where higher values of θ correspond to better performance. She starts with no information, and then observes $n > 1$ signals of the form $s_i = \theta + \epsilon_i$, where the ϵ_i 's are noise terms that are normally distributed with mean 0 and variance σ^2 ; that is, the signals aren't perfect, but are unbiased — correct on average.

If she forms her belief about the politician using Bayes' rule, then the mean of her posterior belief will be the average of all the signals: $\sum_{i=1}^n s_i/n$.⁶ Importantly, when n is large — that is, on issues where she has lots of information — even an extreme value of an individual s_i will have a small impact on her final belief, since the impact is divided by n .

Of course this is a simplistic example, but most of the natural complications only reinforce the conclusion that individual signals are unlikely to matter much. If the citizen does not have an uninformative prior belief, some weight will get placed on the prior belief and less on the individual signals. If the signals are not independent, the informational content of each will be lessened.⁷

Perhaps most importantly, the signals that citizens observe — particularly when studying incendiary misinformation — are rarely unbiased. In the extreme, knowledge of the bias of the sender can completely undermine any learning because the audience can adjust for bias/slant.

To see why, suppose there is a binary state of the world (say, Clinton would be a better president/Trump would be better).⁸ A ‘sender’ knows this state, or at least has some meaningful knowledge that a ‘receiver’ lacks. Assume the receiver knows that the sender wants to persuade her to hold a particular belief — say, that Clinton is the better candidate. The receiver might expect that the sender will always pick the message correspond-

⁴See Kalla and Broockman (2018) for a recent overview as well as a meta-analysis and new evidence that strongly supports the minimal effects hypothesis.

⁵Tellingly, people may be easier to persuade on ‘new’ issues like transgender rights (Broockman and Kalla, 2016).

⁶Of course, the assumption of using Bayes' rule is consequential, frequently criticized in empirical work, and loosened in an increasing volume of theoretical work (Little, 2018). Still, the general principle here — that if there are many messages, the average message can't be too influential — should hold for any updating mechanism.

⁷Two caveats here. First, some recent prominent work indicates that people may not correctly adjust for the correlation in the signals that they receive (Levy and Razin, 2015). Second, absent prior information, a constant correlation between all signals does not change the weight assigned to the individual signals.

⁸What follows is a verbal description of a binary-state version of Crawford and Sobel (1982).

ing to his bias (“Clinton is better”), in which case the message is completely uninformative. And this receiver conjecture is reasonable: if she were to listen to the sender at all, the sender would always want to send the message in favor of his position. So, it can’t be the case that the sender message is informative and the receiver responds to it.

Now, consider a slightly more complex environment where the state variable is continuous (that is, higher values mean an incumbent politician is doing better), and the sender message varies continuously along this dimension.⁹ Set up in a standard way, a typical equilibrium in a model like this involves a sender — who wants the receiver to think the politician is doing well but pays a cost to lie — exaggerating the truth by a fixed amount. To put numbers to it, imagine the content of the message is “On a scale from 1 – 10, how good of a job is Trump doing?”; an equilibrium to this model might involve the sender adding, say, 3 points to what he really thinks. If this is the sender strategy, then the optimal receiver belief is to take what the sender says and subtract 3. The sender can be caught in a trap of expectations: if lying or exaggerating less than expected, the receiver will think Trump is doing *worse* than he really is, and so the sender has to tell costly lies to keep up.

So, the sender’s attempts to persuade run into two problems: the receiver discounts what he says, and if there are other messages out there each sender message has a small impact in the first place.¹⁰

These arguments do suggest some reasons to think fake news might be more dangerous than more traditional political advertising. We have assumed that the receiver knows the bias of the sender, which may not hold in this context. Outlandish made-up stories about Clinton did not end with the disclaimer “I am a Moldovan teenage-troll and I made up this message.” However, there is evidence people are generally able to discern untruthful political claims (Woon, 2017). And Little and Nasser (2018) show that persuasion is difficult even if the receiver is partially ‘credulous’ at the outset; that is, he thinks there is some chance the sender

is honest or unbiased. Further, despite the decisive pro-Trump bias in fake news spread on social media, Trump did better among groups that used the Internet less (Boxell, Gentzkow and Shapiro, 2018).

If we are to be concerned about the effect of misinformation on people’s beliefs, we need reason to think that they actually *believe* the misinformation. However, a general theme from empirical and theoretical work on persuasion — in the U.S. and more comparatively — is that persuasion is really hard.

Another reasonable concern is that even if individual fake news stories have marginal effect, it is possible that heavier exposure to information with a particular skew may matter. For example, several well-identified studies of exposure to partisan television and advertising find an effect on voter behavior in the U.S. (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Spenkuch and Toniatti, 2018). However, the comparative evidence on this point across countries and media technologies is mixed (Enikolopov, Petrova and Zhuravskaya, 2011; Kern and Hainmueller, 2009).¹¹ Still, the most rigorous analysis of fake news to date, discussed in more detail below, indicates that the volume of exposure to fake news is far more modest than in traditional outlets, and as a result likely has a much smaller effect than that identified in these studies (Allcott and Gentzkow, 2017).

Further, when dealing with polarizing issues where in the end people make a binary decision, such as who to vote for, it is unlikely that many are close to changing their minds (Kalla and Broockman, 2018). To connect to the previous example, if there is some critical belief about the politician performance where the citizen votes to re-elect if and only if the average belief about the politician performance is above $\hat{\theta}$, individual signals will only ‘matter’ if they push the belief above or below this threshold. If citizens are very polarized, there will be very few close to this threshold, so even

⁹A classic reference here is Holmstrom (1999); see Little (2015) for a comparative politics application to fraud in authoritarian elections. These models typically assume that the sender does not know the truth and chooses a degree of manipulation to the signal, but under some parametric assumptions the equilibrium is the same if the sender knows the truth (Little and Nasser, 2018).

¹⁰In addition to this ‘equilibrium discounting,’ people are reasonably good at detecting political lies (Woon, 2017).

¹¹See Little (2016) for an overview and theoretical argument for why this might be the case.

¹²Of course, we may be concerned about extremists become more extreme for other reasons; see Stone (2017) for a recent model where

reasonably persuasive arguments may not be decision-relevant.¹² There is good reason to think this observation is highly relevant for fake news, as extreme partisans show the strongest appetite for misinformation (Guess, Nyhan and Reifler, 2018).

Do people believe what they hear? What they say?

An obvious counter to the observations above is that surveys indicate many people *do* believe frequently purveyed falsehoods. Here is a colorful example. A recent *Washington Post* survey showed pictures of the crowds from the Trump and Obama inaugurations, with the latter clearly much larger (Schaffner and Luks, 2017). Despite the ubiquity of these photos in the news, when asked which inauguration corresponded to the photo with the larger crowd, around 40% of Trump voters answered incorrectly (compared to about 10% of Clinton voters and 20% of nonvoters). More striking, when asked the seemingly obvious and factual question of which crowd was larger, 15% of Trump voters answered incorrectly;¹³ presumably those who truly knew which picture was which and wanted to cheerlead their candidate. Beyond the arguably overblown debate over crowd sizes, an Ipsos/Buzzfeed survey found that a majority of all Trump and Clinton voters who recalled seeing false headlines believed them to be somewhat or very accurate (Silverman and Singer-Vine, 2016).

Of course it is easy to find similar examples outside of the United States. A substantial majority of Russians followed their government in blaming Ukrainian forces for the downing of flight MH17 despite strong international consensus to the contrary (Toal and O'Loughlin, 2017). In many countries throughout the world, more people think the United States or Israel were behind the 9/11 attacks than Al-Qaeda, an impression some governments are eager to support, or at least not oppose (Klein, 2008).

But there are good theoretical and empirical reasons to avoid taking these responses too seriously. Both in answering survey questions and in 'real life', people face incentives to engage in preference falsification (Kuran, 1997). While these pressures are likely stronger in more authoritarian settings, Kuran (1997) also includes examples of how answering survey questions on

sensitive topics can lead to conformist incentives outweighing the desire to tell the truth. And when people are given incentives to answer questions correctly, the disagreement among partisans diminishes substantially (Bullock et al., 2015).

Why do such conformist motives lead people to go along with exaggerations or outright falsehoods espoused by elites? Little (2017) suggests one reason. If the less informed members of a group aren't as good at recognizing lies, they will tend to credulously accept what they are told. More informed members of the group recognize what are lies and what aren't, but the pressure to conform makes them go along with the less informed. More concretely, even if a sizable majority of a political party endorses a falsehood, many may realize that what they are saying is unlikely to be true but are simply going along with what they think they are supposed to say on the subject (Bailey, 2017).

More specifically, in the context of the 2016 U.S. election, the most careful studies are not so dire. First, Allcott and Gentzkow (2017) find that while respondents recall (and sometimes believe) prominent fake news headlines, they only do so at a marginally higher rate than 'placebo' stories that the authors made up. So, even if fake news was widely read, it did not seem to leave a lasting impression. Second, only a small slice of the population had heavy exposure to fake news in the first place (Guess, Nyhan and Reifler, 2018); and as discussed above, these are precisely the groups least likely to have their views (and voting decisions) swayed by small or even large amounts of new information.

Persuasion may not be the goal. If persuasion via manipulating information is so hard and has limited success, why would anyone bother doing it in the first place? First, as discussed above, a common theme of models of communication is that if the audience expects to hear from a sender and expects they will exaggerate, the sender may be compelled to send biased messages to keep up with expectations (Holmstrom, 1999; Little, 2015). Even if the audience would be persuaded by only slightly exaggerated stories, those who are free to manipulate information easily may not be able to restrain themselves to tell believable lies (Little

such polarization can occur from common signals and leads to bad outcomes. Svobik (2018) argues that such polarization can lead to tolerance for anti-democratic behavior against political foes. Still, whether fake news as it exists now is really leading to such polarization is an open question (Lazer et al., 2018).

¹³A much smaller fraction of Clinton voters and non-voters made the same error.

and Nasser, 2018)

In the case of leaders, the purpose of spreading misinformation may not be to persuade anyone in the first place. Particularly in authoritarian contexts, getting subordinates to go along with and even repeat crazy lies about leaders' ability can serve as a screening device for loyalty (Marquez, Forthcoming; Crabtree, Kern and Siegel, 2018). While democratic leaders have less to fear from disloyal subordinates, they may use extreme lies for the same purpose (Cowen, 2017).

What about fake news spread by others? A general consensus holds that, in addition to helping Donald Trump, a goal of the Russian intelligence operations during the 2016 elections and beyond was to push contentious issues like racial bias in policing and whether vaccines are safe (Stewart, Arif and Starbird, 2018; Broniatowski et al., Forthcoming). Whether this succeeded in raising the prominence of these debates relative to a counterfactual world without these trolling campaigns or had deeper negative effects on political discourse is an open question.

Other purveyors of false information seem to have a simpler motive: making money. The aforementioned Macedonian teenagers who made thousands of dollars a month getting people (mostly Americans) to read articles touting the imminent indictment of Hillary Clinton likely cared little if the readers believed the story. Their goal was "to generate traffic [by getting] their politics stories to spread on Facebook", and found through trial and error that "the best way to generate shares on Facebook is to publish sensationalist and often false content that caters to Trump supporters." An American creator of fake news tracked down by NPR reached the same conclusion (Sydell, 2016). And while he claimed his goal was to discredit the extreme right by exposing their gullibility to false stories, he also admitted that making ten to thirty thousand dollars a month was an incentive as well.

Millions of people saw the famous Onion headline "Drugs win Drug War", and the advertising revenue from viewers of this and other satirical writings is how "America's Finest News Source" makes money. But this headline did not lead to citizens rising up and calling for the resignation of top officials at the Drug Enforcement

Agency as punishment for their failure. Perhaps the effect of more malicious-seeming fake news is more similar to this open satire than the death-of-truth brigade would have you think.

III. All It Takes Is A Little Responsiveness

In sum, we can't infer from the prevalence of fake news and the like that opinions are being meaningfully changed or society further polarized. We simply don't have any direct evidence of fake news having a large impact on political beliefs or behavior (Lazer et al., 2018). While it is possible that future research will uncover such evidence, what we know about strategic communication and information manipulation in general does not make this seem likely.

I admit it is hard to completely shrug off advisors to the most powerful person in the world insisting on his right to "alternative facts" and the fact that most of his allied politicians and voters seem unbothered by this rhetoric. But even if one doubts the arguments above, holding politicians accountable for their performance doesn't require every citizen to digest unbiased signals in an objective and rational manner.¹⁴ As long as we resolve political differences with fair elections, even a modest number of centrists who generally respond to how their politicians behave in a reasonable manner (particularly paired with an energized opposition) can defeat incumbents with completely ossified and information-resistant bases.

References

- Allcott, Hunt and Matthew Gentzkow. 2017. "Social media and fake news in the 2016 election." *Journal of Economic Perspectives* 31(2):211–36.
- Ashworth, Scott and Ethan Bueno de Mesquita. 2014. "Is voter competence good for voters?: Information, rationality, and democratic performance." *American Political Science Review* 108(3):565–587.
- Bailey, Ronald. 2017. "Trump, propaganda, and credulity." *Reason*, February 24.
- Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker and Richard Bonneau. 2015. "Tweeting from left to right: Is online political communication more than an echo chamber?" *Psychological Science* 26(10):1531–1542.
- Berelson, Bernard R., Paul F. Lazarsfeld and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in*

¹⁴In fact, Ashworth and Bueno de Mesquita (2014) show that deviations from standard rationality can improve voter welfare by changing the behavior of politicians.

- a *Presidential Campaign*. Chicago: University of Chicago Press.
- Boxell, Levi, Matthew Gentzkow and Jesse M. Shapiro. 2018. "A note on internet use and the 2016 US presidential election outcome." *PLOS ONE* 13(7).
- Broniatowski, David A., Amelia M. Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C. Quinn and Mark Dredze. Forthcoming. "Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate." *American Journal of Public Health*.
- Broockman, David and Joshua Kalla. 2016. "Durably reducing transphobia: A field experiment on door-to-door canvassing." *Science* 352(6282):220–224.
- Bullock, John G., Alan S. Gerber, Seth J. Hill and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10(4):519–578.
- Cowen, Tyler. 2017. "Why Trump's staff is lying: It's both a loyalty test and a proclamation of power." *Bloomberg*, January 23.
- Crabtree, Charles, Holger L. Kern and David A. Siegel. 2018. "Cults of personality, preference falsification, and the dictator's dilemma." Unpublished manuscript, University of Michigan.
- Crawford, Vincent P. and Joel Sobel. 1982. "Strategic information transmission." *Econometrica: Journal of the Econometric Society* 50(6):1431–1451.
- DellaVigna, Stefano and Ethan Kaplan. 2007. "The Fox News effect: Media bias and voting." *Quarterly Journal of Economics* 122(3):1187–1234.
- Enikolopov, Ruben, Maria Petrova and Ekaterina Zhuravskaya. 2011. "Media and political persuasion: Evidence from Russia." *American Economic Review* 101(7):3253–3285.
- Flaxman, Seth, Sharad Goel and Justin M. Rao. 2016. "Filter bubbles, echo chambers, and online news consumption." *Public Opinion Quarterly* 80(S1):298–320.
- Gehlbach, Scott, Konstantin Sonin and Milan Svobik. 2016. "Formal models of nondemocratic politics." *Annual Review of Political Science* 19:565–584.
- Guess, Andrew, Brendan Nyhan and Jason Reifler. 2018. "Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign." Unpublished manuscript, Princeton University.
- Guriev, Sergei M. and Daniel Treisman. 2015. "How modern dictators survive: Cooptation, censorship, propaganda, and repression." CEPR Discussion Paper No. DP10454.
- Holmstrom, Bengt. 1999. "Managerial incentive problems: A dynamic perspective." *Review of Economic Studies* 66(1):169–182.
- Horz, Carlo. 2018. Informational and non-informational propaganda. In *Fake News and the Politics of Misinformation, Comparative Politics Newsletter*, ed. Matt Golder and Sona Golder. Vol. 28(2).
- Kalla, Joshua L. and David E. Broockman. 2018. "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." *American Political Science Review* 112(1):148–166.
- Kern, Holger Lutz and Jens Hainmueller. 2009. "Opium for the masses: How foreign media can stabilize authoritarian regimes." *Political Analysis* 17(4):377–399.
- Kessler, Glenn, Salvador Rizzo and Meg Kelly. 2018. "President Trump has made 3, 251 false or misleading claims in 497 days." *Washington Post*, June 1.
- Klein, Kent. 2008. "Global poll shows doubt about al-Qaida role in 9/11 attacks." *VOA News*, September 10.
- Kuran, Timur. 1997. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge: Harvard University Press.
- Lazer, David M.J., Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild et al. 2018. "The science of fake news." *Science* 359(6380):1094–1096.
- Levy, Gilat and Ronny Razin. 2015. "Correlation neglect, voting behavior, and information aggregation." *American Economic Review* 105(4):1634–1645.
- Little, Andrew T. 2015. "Fraud and monitoring in non-competitive elections." *Political Science Research and Methods* 3(1):21–41.
- Little, Andrew T. 2016. "Communication technology and protest." *Journal of Politics* 78(1):152–166.
- Little, Andrew T. 2017. "Propaganda and credulity." *Games and Economic Behavior* 102:224–232.
- Little, Andrew T. 2018. "The distortion of related beliefs." Unpublished manuscript, University of California, Berkeley.
- Little, Andrew T. and Sherif Nasser. 2018. "Unbelievable lies." Unpublished manuscript, University of California, Berkeley.
- Lozada, Carlos. 2018. "Can truth survive this president: An honest investigation." *Washington Post*, July 13.
- Marquez, Xavier. Forthcoming. "Two models of political leader cults: Propaganda and ritual." *Politics, Religion, & Ideology*.
- Martin, Gregory J. and Ali Yurukoglu. 2017. "Bias in cable news: Persuasion and polarization." *American Economic Review* 107(9):2565–99.
- Nyhan, Brendan. 2018. "Fake news and bots may be worrisome, but their political power is overblown." *New York Times*, February 13.
- Plante, Chris. 2016. "This Is Fine creator explains the timelessness of his meme." *The Verge*, May 5.
- Rich, Michael D. and Jennifer Kavanagh. 2018. *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. Rand Corporation.
- Schaffner, Brian and Samantha Luks. 2017. "This is what Trump voters said when asked to compare his inauguration crowd with Obama's." *Washington Post*, January 25.

- Silverman, Craig and Jeremy Singer-Vine. 2016. "Most Americans who see fake news believe it, new survey says." *BuzzFeed News*, December 6.
- Spenkuch, Jörg L and David Toniatti. 2018. "Political advertising and election results." *Quarterly Journal of Economics* (133):1981–2036.
- Stewart, Leo G., Ahmer Arif and Kate Starbird. 2018. Examining trolls and polarization with a retweet network. In *Proc. ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web*.
- Stone, Daniel F. 2017. "Just a big misunderstanding? Bias and Bayesian affective polarization." Unpublished manuscript, Bowdoin College.
- Sunstein, Cass R. 2018. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press.
- Svolik, Milan W. 2018. "When polarization trumps civic virtue: Partisan conflict and the subversion of democracy by incumbents." Unpublished manuscript, Yale University.
- Sydell, Laura. 2016. "We tracked down a fake-news creator in the suburbs. Here's what we learned." National Public Radio, November 23.
- Toal, Gerard and John O'Loughlin. 2017. "Why did MH17 crash?: Blame attribution, television news and public opinion in southeastern Ukraine, Crimea and the de facto states of Abkhazia, South Ossetia, and Transnistria." *Geopolitics* pp. 1–35.
- Woon, Jonathan. 2017. "Political lie detection." Unpublished manuscript, University of Pittsburgh.

Why Do Governments Censor? Expanding from State Survival to Content Regulation Theories in Political Science

by Stephen Meserve
Texas Tech University

The utopian technologist conceptual frame of the Internet as a free-for-all platform, fundamentally beyond private or public control, is gone. Digital content generally passes through one or more private points of control (Zittrain, 2003), whether this is an Internet service provider such as Comcast, a digital service provider such as Google, or a social media network such as Twitter. Content control and management is receiving increased public scrutiny due to controversies associated with the 2016 U.S. presidential election, combined with

more general anxieties about how content serving algorithms affect individual political attitudes, democratic discourse, and political behavior (Allcott and Gentzkow, 2017; Sunstein, 2017). But when democratic publics discuss how access to digital content should be restricted, those discussions are usually framed in terms of how private companies should manage their own platforms rather than how governments should manage content on the Internet.

It is sometimes taken for granted that governments themselves are largely hands-off with respect to digital content in democracies. Even within political science, most of the empirical investigations into content restriction and censorship take place in the context of authoritarian regimes. An increasingly rich literature focuses on the strategies and motivations of authoritarian regimes with respect to digital content (Roberts, 2018) and the ways in which protest movements and coercive regimes react to one another in the digital context (Tufekci, 2017). The academic discussion is much more muted, however, and different in tone, when it turns to democracies, and especially developed democracies, usually focusing on elections or the polarization of the public (Farrell, 2012). Yet, democracies engage in a great deal of content removal. By and large, the difference between democracies and authoritarian regimes is that content removal in democracies is done indirectly through private points of control — leaning on private actors to censor digital materials and prevent public access. The contrast in strategy between democracies and authoritarian regimes in the digital realm is quite stark. Authoritarian regimes tend to rely on more direct means of content removal and restriction, either setting up control systems themselves or outright restricting companies that might publish or link to objectionable content from doing local business, while democracies typically take down content from the Internet using legal institutions or through executive or police requests to private points of control.¹

From the perspective of empirical social science, one of the most intriguing elements of democratic use of private points of control for censorship activities is that they generate traceable data that is broadly comparable across countries. Private points of control like

¹The recently publicized Google search engine created to comply with Chinese censorship laws, Dragonfly, is a relatively novel way of approaching authoritarian digital content restriction (Gallagher, 2018). Google had previously been restricted from operating within China by the government via more direct means. Similarly, indirect means through private points of control are not the only way democracies restrict the Internet, just the most typical and visible way they do so. Other methods of content control taken by national security institutions in democracies, for example, sometimes do not fit into this simplified narrative.