

THE DEMON OF THE BELFRY

INVESTIGATING THE EVOLUTION OF THE PENNY PRESS AND THE EMMANUEL CHURCH MURDERS

ALSO FEATURING

- The Digital Pandemic
- The Force on the Front



THE DIGITAL PANDEMIC

Averting Negative Externalities of Disinformation Through Platforms

RYAN CATTICH

IN RESPONSE TO VIRAL SPREADS OF DISINFORMATION IN THE 21ST CENTURY, MOST NOTABLY IN 2020, THIS PAPER ANALYZES SECTION 230 OF THE COMMUNICATIONS DECENCY ACT AS CURRENTLY WRITTEN. PRESENTLY, MUCH AMBIGUITY SURROUNDS THE ACTUAL PURPOSE AND IMPLICATIONS SECTION 230 DENOTES FOR THE UNITED STATES, AND THE WORLD. HEADS OF THE LARGEST TECHNOLOGY COMPANIES FREQUENTLY COME UNDER FIRE, CALLED UPON BY BOTH POLITICIANS AND THE PUBLIC TO COMBAT THE SPREAD OF DISINFORMATION ON THE INTERNET. IN THIS PAPER, I PRESENT AN ANALYSIS OF SECTION 230 FROM ITS INCEPTION TO TODAY, FOLLOWED BY AN EXAMINATION OF PLATFORM-BY-PLATFORM DECISIONS ON HOW TO BEST COMPLY WITH THE STATUTE. I ARGUE REVOKING SECTION 230 IS NOT A REASONABLE SOLUTION, RATHER A STATUTORY MODIFICATION COUPLED WITH HEIGHTENED INDUSTRY COLLABORATION WILL SERVE TO UPHOLD THE SECTION'S ORIGINAL OBJECTIVES SET FORTH IN 1996.

"This is the cost of disinformation" – Atul Gawande, responding to nurse Jodi Doering's story of her COVID-19 patients' denial of their condition1

I. INTRODUCTION

Eight months into a global pandemic and it appears the spread of the novel coronavirus is supplemented by the spread of disinformation. One story, in particular, illustrates just that – South Dakota ER nurse Jodi Doering recounts a few days' worth of COVID-19 patients in a Twitter thread: "The ones that stick out are those who still don't believe the virus is real...They tell you there must be another reason they are sick. They call you names and ask why you have to wear all that 'stuff' because they don't have COVID because it's not real" (Orth, 2020). How did we get here?

Doering's story draws a destitute parallel to developing countries. Despite effective remedies like mosquito nets and deworming pills, people in developing countries neglect to use them at large, partly citing a lack of available information, or misinformation more generally (Pulford et al., 2011). In the United States, with an abundance of readily accessible information at our fingertips, how are we, too, in a position where people neglect reality? Outside of the externalities inherent in a pandemic, widespread disinformation hinders and undermines the very integrity and the accuracy of fact. Disinformation has the potential to influence civic engagement, election results, and, most unfortunately, downstream health outcomes. In search of a remedy, we must begin at the root of the problem: section 230 of the Communications Decency Act.

In so doing, the evaluations and recommendations in this paper contribute to a literature on the effects of media and disinformation on behavior and health outcomes (Bursztyn et al., 2020; Banerjee et al., 2019a; Bursztyn et al., 2019; La Ferrara, 2016; DellaVigna and La Ferrara, 2015; La Ferrara et al., 2012; Chiang and Knight, 2011; Jensen and Oster, 2009). Prior work has revealed that media exposure can increase hate crimes (Muller and Schwarz, 2018; Bursztyn et al., 2019) and mass killings (Yanagizawa-Drott, 2014); it can also affect domestic violence (Card and Dahl, 2011; Banerjee et al., 2019b), fertility choices (La Ferrara et al., 2012; Kearny and Levine, 2015), and responses to natural disasters (Long et al., 2019).

Related to this analysis is contemporaneous work calling for a reinterpretation and a reassessment of section 230 of the Communications Decency Act as it stands today. Lotty (2020) argues the current interpretations of the scope of section 230 immunity wrongfully deny individuals who have been sexually harassed or assaulted an opportunity to hold online services accountable for causing or exacerbating their harms (Lotty, 2020). Thus, Lotty prescribes a necessary revision to section 230 to better align with contemporary views regarding the role of the internet user and the responsibility of technology companies to deter sexual misconduct. In a similar vein, Sloss (2020) responds to the Myanmar military carrying out brutal attacks via Facebook against Rohingya Muslim communities in Rakhine State in 2017. Going further than Lotty, Sloss argues for a statutory exception in 230 to permit civil suits against internet companies on account of their alleged complicity in genocide, war crimes, or crimes against humanity (Sloss, 2020). With a civil liability exception, Rohingya plaintiffs could then bring a state tort law claim against Facebook alleging that Facebook was negligent in permitting its platform to be utilized to spark mass violence against the Rohingya in Myanmar. As the law stands today, such a case would be immediately nullified under the current federal preemption defense to state tort law claims.

Irrespective of the study, the written law, or the lack there-of, section 230 most certainly demands a statutory modification – a modification which provides a positive incentive for companies to internalize the risk prevention of disseminating misinformation on their platforms. The remainder of this paper proceeds as follows. In section 2, I provide a brief history of the Communications Decency Act, followed by a substantive overview of the clauses under section 230. In section 3, I present a platform-by-platform analysis of their prior and current efforts to prevent the spread of disinformation. In section 4, I introduce potential amendments to the current legislation, present notable downstream implications, and provide empirical evidence of disinformation as a negative externality to society. Section 5 concludes.

2. HISTORY

2.1 Inception of The Communications Decency Act

Introduced in February of 1995 by Senator James Exon, the federal Communications Decency Act (CDA) was created to combat a growing issue of extensive pornography and obscenity on the internet (Cannon, 1996). As passed, the CDA extends the "antiharassment, indecency, and antiobscenity" restrictions currently placed on telephones to

interactive computer service providers ("ICSPs"). The bill was promptly met with opposition from some lawmakers and interest groups who opposed the idea of interfering with the Internet, along with concerns that the act was a "violation of free speech and ... of the right of adults to communicate with each other" (Cannon, 1996). Opposing the CDA on the basis of these concerns, coupled with the 1995 New York Supreme Court decision *Stratton Oakmont v. Prodigy Services Co.*, representatives Christopher Cox and Ron Wyden proposed the "Cox-Wyden Amendment" to Exon's bill, some parts of which would ultimately become section 230 of the CDA as it stands today.

In Stratton Oakmont, the court held that an online bulletin board service provider could be held liable for defamatory content posted by users on its platform, given that the website proactively monitored, screened, and removed offensive user content, thereby serving as an editor and publisher of all posted content and assuming legal responsibility for it (Stratton Oakmont v. Prodigy Services Co., 2018). By contrast, a 1991 New York case, Cubby, Inc. v. CompuServe, Inc., held that an ICSP that did not regulate third-party user content evaded liability for libel since it did not know of and had no editorial control over posted defamatory material. Astonishingly, Stratton Oakmont asserted that an online service that does nothing by means of monitoring for problematic content can never be legally responsible for the content of its users; in addition, a service that takes good-faith steps to screen such content subjects itself to liability (Ardia, 2010).

As a direct response, the Cox-Wyden Amendment was proposed to encourage online services to take proactive measures to improve online safety and regulate objectionable content without fear of increased liability (Ardia, 2010). Thus, this amendment allowed private ICSPs to address

the problem of online indecency, while concurrently upholding the Representatives' policy goal of fostering the "vibrant and competitive free market that presently exists for the Internet and other interactive services" (Stratton Oakmont v. Prodigy Services Co., 2018). Subsequently, Congress passed the Telecommunications Act of 1996, which included the Communications Decency Act and the Cox-Wyden Amendment – legislation presently regarded as section 230 of the CDA (Ardia, 2010).

2.2 Section 230 of the Communications Decency Act

Section 230 of the Communications Decency Act is the most significant piece of legislation regarding internet regulation, praised by commentators for allowing the Internet as we know it today to prosper. This statute largely restricts the scope of potential liability for interactive computer service providers based on content transmitted by third-party users of the service and an ICSP's good-faith actions to limit access to objectionable material transmitted by users. The ICSP definition extends to "any information service, system, or access software provider that provides or enables computer access by multiple users to a computer service," which includes most websites and Internet-based apps (Stratton Oakmont v. Prodigy Services Co., 2018).

As written, section 230 bears the name "Protection for private blocking and screening of offensive material" (Stratton Oakmont v. Prodigy Services Co., 2018). Section 230(b) lists the policy objectives, which include: promoting the continued development of the Internet, preserving the free market that is the Internet, spurring technological innovation that maximizes user control, eliminating "disincentives for the development and utilization of blocking and filtering technologies," and guaranteeing the enforcement

"[Section 230 of the Communications Decency Act] largely restricts the scope of potential liability for interactive computer service providers based on content transmitted by third-party users of the service and an ICSP's good faith actions to limit access to objectionable material transmitted by users."

22

22

of federal online security laws (Stratton Oakmont v. Prodigy Services Co., 2018). Section 230(c) sheds light on what constitutes "Good Samaritan" neutral actors and publishers online. Below the heading "Protection for 'Good Samaritan' blocking and screening of offensive material," section 230(c)(1) states: "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" (Stratton Oakmont v. Prodigy Services Co., 2018).

Successively, section 230(c) safeguards civil liability for any "provider or user of an interactive computer service" given "any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable," irrespective of if such material is protected by the constitution (Stratton Oakmont v. Prodigy Services Co., 2018).

Overruling *Stratton Oakmont*, this legislation issues ICSPs that make bona fide efforts to monitor and screen content immunity from legal actions based on user content generated or transmitted on their platforms. Nevertheless, the Act explicitly does not permit ICSP immunity from liability for specific offenses: violations of federal criminal law, communications privacy law, intellectual property law, and federal sex trafficking law (Stratton Oakmont v. Prodigy Services Co., 2018).

However, not included in the legislation: what constitutes such bona fide voluntary action, which explicit actions separate platforms from publishers, and how these specific, yet ambiguous policy goals can be *efficiently* achieved. At present, courts are progressively interpreting section 230 to broadly restrict the scope of potential civil and state law liability for ICSPs, thus fostering a safe environment for web-based corporations against a multitude of claims. How those companies do so is, evidently, up to them.

3. PLATFORM-BY-PLATFORM EVALUATION

As of now, technology companies and media platforms are left to self-regulate, experimenting on their own with various moderation strategies and models of platform governance. Essentially, individual companies are tasked with not only aligning their platform objectives with the needs of their users but to promote the principles of the common good as well.

3.1 Twitter

In an interview with Boston College Magazine, Colin Crowell, former vice president of global public policy and corporate philanthropy at Twitter, discusses the ramifications of Twitter flagging one of Donald Trump's tweets for the first time (Figure 1). Signaling a new approach to combatting disinformation, Crowell asserts that "navigating the emerging and evolving online terrain of disinformation to protect the integrity of vital civic conversations, without succumbing to excessive censorship, is critical to safeguarding the internet as a vibrant platform for human expression" (Tempera, 2020). Here, Crowell echoes Twitter CEO Jack Dorsey's overarching commitment to serving the public conversation. Although Twitter does not have an explicit policy on general disinformation, Twitter does present specific guidelines and policies on the following areas: violent threats policy, glorification of violence policy, platform manipulation, synthetic and manipulated media policy, and, most notably, civic integrity policy.

According to Crowell, "when you have a platform as large as Twitter, it is important to recognize when there is a need to help users understand what they are viewing with additional context," and Twitter's civic integrity policy does just that (Tempera, 2020). Updated in October of 2020 ahead of the presidential election, Twitter asserts they will "label or remove false or misleading information about how to participate in an election or other civic process" (Twitter, 2020). This policy includes but is not limited to posting or sharing content that may suppress participation or mislead people about when, where, or how to participate in a civic process. Further, Twitter explicitly commits to removing and labeling false or misleading information intended to intimidate or dissuade people from participating in a civic process, in conjunction with any intent to undermine public confidence in an election. On the same webpage, Twitter includes the subheadings: "What is not a violation of this policy?"; "Who can report violations of this policy?"; and "What happens if you violate this policy?" (Twitter, 2020). Thus, Twitter makes it abundantly clear both the ramifications and the qualifications of their policy violations in an effort to transparently combat negative externalities to society. But is Twitter doing to prevent the spread of disinformation with this approach?

In defense of Twitter's civic integrity policies, Crowell praises the decision to take action. The flagging feature, according to Crowell, "provide[s] a screen that [says], basi-

cally, this tweet is in violation of Twitter rules, but it is remaining on the service in the public interest," as figures like Donald Trump are democratically elected heads of state (Tempera, 2020). Firmly, Crowell asserts that "Twitter can and should take action when online speech risks offline harm," touching on negative societal externalities like disenfranchising voters and downplaying risks of CO-VID-19 (Tempera, 2020). Previously, Twitter only had the binary choice of either removing a tweet or leaving it up; now, Twitter allows journalists and the public to comment, links access to factual information, and takes steps to limit the ability of content to reach viral spread (Tempera, 2020). In accordance with section 230, simply flagging tweets does not invoke publisher status on Twitter's behalf; however, Twitter begins to stray from the comfort of the "neutral actor" position.² The existence of such policy pages represents an act in good faith on part of Twitter to screen for objectionable content, thus shielding them from civil and state liabilities. Nevertheless, Twitter is well within its right as an ICSP to prohibit attempts to use its services to manipulate or disrupt civic processes, along with distributing false or misleading information.

3.2 Facebook

In a testimony before the United States Senate Committee on the Judiciary, Facebook CEO Mark Zuckerberg verbalizes Facebook's mission to "'give people the power to build community and bring the world closer together" (Testimony of Mark Zuckerberg Facebook, Inc., 2020). Zuckerberg touts that Facebook removed false claims about civic processes and displayed more than 150 million warnings on posts, after review by independent third-party factcheckers (Testimony of Mark Zuckerberg Facebook, Inc., 2020). However, in response to Twitter's fact-check labels, Zuckerberg takes a different approach, "[believing] strongly that Facebook shouldn't be the arbiter of truth," in an interview with Fox News (Halon, 2020). Nevertheless, Facebook, too, makes an explicit commitment to stop the spread of misinformation.

Under the heading, "Working to Stop Misinformation and False News," Facebook makes its objectives clear in fostering accurate information on its platform. Specifically, focus is given to three areas: disrupting economic incentives, building new products, and helping people make more informed decisions (Facebook, 2020). Citing financial motivations behind fake news, Facebook commits to removing the economic incentives for traffickers of misinformation - doing so through third-party fact-checking organizations, strict enforcement of their policies, and applying machine learning to assist response teams (Facebook, 2020). Reinforcing Zuckerberg's stance against becoming arbiters of truth, Facebook is "building, testing and iterating on new products to identify and limit the spread of false news," engaging with their community "Oversight Board" to elicit help in fact-checking (Facebook, 2020). In doing so, Facebook is explicitly making bona fide efforts to screen for misinformation, upholding an impartial position, rather than a publisher position in working with external, unbiased groups. Through the "News Integrity Initiative," Facebook is pulling together over 25 funders and participants, including tech industry leaders, academic institutions, non-profits, and third-party organizations, to help people make informed judgments about what they read online.35 As a result of this initiative, Facebook positions itself to advance global news literacy, bolster trust in journalism, and better inform the public conversation.

Irrespective of hate speech and promotion of violence violations, Facebook takes a hands-off approach to content moderation. Per the "Community Standards" page, Facebook makes its position clear: "We want to help people stay informed without suppressing productive public debate," justifying the rationale behind not removing fake news from the Platform (Facebook, 2020). Rather, Facebook makes good faith efforts to reduce the circulation of misinformation by adjusting their algorithms so as to "display [fake news] further down in the news section," thus averting viral spread (Facebook, 2020). Essentially, Facebook encourages users to decide for themselves what to read,

"Voluntary action superseded by profit in no way advances the effort on combatting the spread of misinformation; thus, there must be a disincentive to the 1 step forward, 2 steps back approach."

what sources to trust, and what content to post, all the while attempting to not overstep their boundaries as a neutral actor.

3.3 YouTube

In the time leading up to the 2020 presidential election, a YouTube ad likening Democratic nominee Joe Biden to Venezuelan socialism circulated in the state of Florida (Donald J. Trump, 2020). Despite a published fact-check by the Associated Press, YouTube showed the ad more than 100,000 times in the eight days leading up to the election (Merrill & McCarthy, 2020). Trump went on to win the state of Florida by roughly 375,000 votes – the largest margin in a presidential election there since 1988 – and carried about 55% of the Cuban American vote (Merrill & McCarthy, 2020). In retrospect, this ad illustrates key gaps in the policing of misinformation by Google, the parent company of YouTube; while nominally prohibiting false claims in advertising, Google rarely takes down political ads (Google, 2020).

However, YouTube takes a more nuanced approach to this issue. Per the "How does YouTube combat misinformation" page, YouTube engages a range of practices: removing content in violation of policies, raising up authoritative sources for news and information, and reducing recommendations of borderline content and harmful misinformation (YouTube, 2020). In an attempt to satisfy the voluntary action conditions of section 230, algorithms go as far as to not proactively recommend content that comes close to violating Community Guidelines; thus, taking action to limit the spread of disinformation. To determine what qualifies as "harmful" misinformation, YouTube states that they consult "external human evaluators" and experts to determine whether content promotes unsubstantiated conspiracy theories, or inaccurate information - reviewing "hundreds of thousands of hours" of video each day (YouTube, 2020). Unarguably, YouTube explicitly makes commitments and takes action to remove objectionable content, within their rights as an ICSP. Despite these articulated commitments, no capacity exists in which YouTube can be held accountable for defaulting on their expressed promises on deterring misinformation.

Circling back to the YouTube ad; if the Associated Press fact-check does not qualify as an external expert evaluation on unsubstantiated conspiracy theories, what does? Negating their own policy, company spokeswoman Charlotte Smith claims that the campaign ad "does not violate our

policies...and we're not going to attempt to adjudicate every claim or counterclaim" on the platform (Merrill & Mc-Carthy, 2020). Albeit not a direct violation of section 230, this event underscores the urgent need for companies to be held accountable for what spreads on their platforms, rather than merely relying on an honor system for efficient platform governance. Given people watch over a billion hours of video any given day on YouTube, financial incentives could be a driving force behind the platform falling short of enforcing its policies ("Press - Youtube", 2020). In 2020 alone, the Trump campaign spent \$106 million -\$37.2 million in the last month of the campaign – on both YouTube and Google search ads alone (Thompson, 2020). Incalculable, however, is the effect running these ads had on disenfranchising and dissuading potential Latin American voters in the state of Florida. With tens of millions of dollars in prospective ad revenue on the line, the clauses currently set forth under section 230 do little to counter any perverse commercial incentives at companies like YouTube. Thus, although YouTube makes voluntary efforts to publish their commitments like Twitter and Facebook, the commitments themselves appear vague and uncorroborated by decisive action.

3.4 Google

In 2019, Google was heavily criticized in a 2019 Senate Investigation for its role in spreading misinformation during the 2016 presidential election (Wells, 2019). In October of that same year, a Senate Intelligence Committee released a report describing just how fast misinformation can spread on the platform; just days after the 2016 election, a false news story claiming that Trump had won the popular vote ranked higher on Google than accurate stories (Cassady, 2020). Although a mixed bag in its history with politics, Google's present-day initiatives to combat the spread of misinformation are centered around fact-checking.

Specifically, the "Google News Initiative" (GNI) is supporting "First Draft," a nonprofit protecting communities from harmful misinformation, with a \$6.5 million grant (Mantzarlis, 2020). With an immediate focus on coronavirus information, the funding goes directly to fact checkers fighting misinformation around the world. In response to misinformation circulating about COVID-19, Google is supporting journalism fellowships at Stanford University to "collate data from around the world and help journalists tell data-driven stories that have [an] impact in their communities" (Mantzarlis, 2020). Additionally, Google says it

"There will always be somewhat of a tradeoff between disappointing people and promoting the common good."

has made Google Trends data available in "localized pages with embeddable visualizations," in an effort to help journalists understand and meet people's information needs around the world. However, despite the promise of these initiatives, they appear reactionary. In March of 2020, the Global Disinformation Index found that 1,400 sites spreading COVID-19 misinformation earned a collective \$76 million in ad revenue, with the vast majority coming from Google (Figure 2). Thus, Google seemingly understands the dilemma: they can take down a conspiracy theorist's YouTube channel for making false claims about CO-VID-19; however, continuing to run ads on the website negates the entire purpose. Voluntary action superseded by profit in no way advances the effort on combatting the spread of misinformation; hence, there must be a disincentive to the "I step forward, 2 steps back" approach.3

4. PROPOSALS AND IMPLICATIONS

Passing judgment on the regulation of free speech entails a balancing act, as Zuckerberg calls it, "of competing equities" (Testimony of Mark Zuckerberg Facebook, Inc., 2020). In some cases, the right thing to do from a safety or security viewpoint is not the most ideal for privacy or free expression; thus, there will always be somewhat of a tradeoff between disappointing people and promoting the common good.4

Deciding upon such tradeoffs is anything but straightforward and passing down this moderation to individual companies raises much ambiguity surrounding the enforcement of section 230. Given the relative ease of switching between platforms, augmentation of the statute is of the uttermost importance - working with technology companies, rather than against them.

4.1 Statutory recommendations

First, Congress should create a statutory exception in section 230(c) that permits civil liability against any ICSP's complicity in the spread and circulation of disinformation. Not uncommon, last year lawmakers passed a bill that added a new exception to section 230 that positions platforms to be liable for any third-party content that facilitates sex trafficking (Romano, 2018).

In addition to modifications in this clause, section 230(c) must unambiguously specify what constitutes "any voluntary action taken in good faith" to provide clear guidance to all ICSPs on the explicit steps they must take (Stratton Oakmont v. Prodigy Services Co., 2018).

Parallel to Sloss (2020) in his push to hold companies accountable for complicity in mass genocide, a statutory exception that withdraws civil liability could also work in internalizing the negative externalities of disinformation. Under section 230(c), such an exception to liability would arise in the case that a company fails to prevent transmission of inaccurate information if that information: a) would be understood by ordinary readers as incitement or inducement to mislead or suppress people of factual information and the written law; and b) there is a significant risk that any recipient of the inaccurate information is exposed to harm. The statute should also include a duty for companies to remove such content within the 24-hour time frame since the content's inception. Nevertheless, the statute should continue to preserve immunity from civil liability for any company that makes a reasonable, bona fide effort to comply with content removal, but is unable to do so for a justifiable reason. To do so, however, such bona fide efforts must be clearly defined in section 230(c).

Between Facebook and Twitter alone, the approaches to taking preventative steps in combatting misinformation are noticeably different. On the one hand, flagging violations and linking them to factual journalism raises concerns of anti-conservative social media bias; however, there is currently no evidence to support this exists (Ingram, 2019). On the other hand, allowing users to simply decide for themselves what to read does little by means of taking tangible, good faith steps in countering false information. Ideally, it is the work of journalists to hold those in power accountable and to provide context to readers; however, given the scale of platforms like Twitter and Facebook, the clauses in section 230 must recognize the need to take action. Thus, to define what proper good faith action is in practice, the legislation should mandate aspects of both the Twitter and Facebook approaches in outlining what such voluntary steps look like for all ICSPs. The approach should: require companies to take action to



limit the ability of disinformation to go viral through algorithmic modifications, flag inaccurate information linked to third-party fact checks, and allow content to be seen such that people can debate it, support it, refute it, or dissect it, and do so publicly.⁵

4.2 Consequential incentives

An exception to section 230 that withdraws liability from companies that perpetuate disinformation could have a preventative effect by providing a positive incentive for companies. Typically, the *mens rea* for civil liability is a negligence standard, rather than specific intent. Such a negligence standard would incentivize companies to adopt sensible measures to inhibit third parties from violating their community guidelines and policies on misinformation. In fact, given the global reach of major platforms like Twitter, Facebook, and Google, these companies arguably have greater power to prevent negative externalities to society than do some national governments.

One could argue that ICSPs do not need additional incentives to limit users from posting inaccurate information because they have already taken steps to address the problem. For example, after hearing about violence in Myanmar, Facebook removed accounts associated with the Myanmar military and hired Burmese language speakers to monitor content (Ellis-Peterson, 2018). However, Facebook did not take such actions until after users in Myanmar already used the platform to spark genocidal violence (Stecklow, 2018). Thus, a federal statute subjecting companies to potential civil liability would incentivize companies to act proactively to prevent harms before they occur, rather than issuing reactionary responses. To avoid liability, companies would need, among other things, to hire teams to specifically monitor objectionable content; hence, the liability risk provides a financial incentive for companies to incur that cost to avoid future adverse judgments. Without such an incentive, ICSPs would neglect to incur the costs of risk prevention until after it is too late.

Additionally, we can leverage aspects of behavioral theory in order to "nudge" platform users into taking their *own*

steps in combatting the spread of disinformation.⁶ In the introduction to their book Nudge, Thaler and Sunstein present a nudge as "any aspect of choice architecture that alters people's behavior in a predictable way without forbidding any Options" (Thaler & Sunstein, 2008). Such an intervention closely mirrors what Twitter already does: flagging tweets in violation of its community guidelines, nudging users to follow the link to accurate information. As Thaler and Sunstein put it, "Putting the fruit at eye level counts as a nudge. Banning junk food does not" (Thaler & Sunstein, 2008). Twitter, through flagging, places accurate information at eye level, rather than banning it altogether. Along with linking accurate information, Twitter can also leverage advancements in machine learning in order to curate feed algorithms in such a way that presents information in an objective, unbiased manner. In totality, such behavioral nudges demand more in the way of freedom of choice and less so in the way of government constraint and compulsion.

4.3 "Revoke 230" is not the solution

As recently as December 2020, President Trump threatened to veto an annual defense bill unless Congress revoked section 230 of the CDA (Brandom, 2020). Just one year prior, in December of 2019, now president-elect Joe Biden, in an interview with *The New York Times*, said "'section 230 should be revoked, immediately should be revoked, number one'" (The New York Times Editorial Board, 2020). Given the Communications Decency Act is now over 20 years old, lawmakers are trying to amend the statute to best align with present-day needs; simply put, all-out revoking section 230 is not the solution.

Above all, section 230 serves one main purpose: protecting ICSPs from the liability for the speech of others on their platforms. If it were to be revoked, former Twitter VP Crowell suggests doing so "creates a dilemma for the companies because they would have to over censor and start taking down anything that might remotely run a liability risk," noting it would cause a ripple effect in the timeliness of the internet (Tempera, 2020). Thus, before any content could be posted on an ICSP, it would essentially

"Revise, revoke, repeal — whatever 'R' word you choose — the debate over the future of section 230 boils down to one, simple idea: the law exists to protect people."

trajectories in the broader population (Bursztyn, 2020).

have to be approved by lawyers, having the involuntary effect of making the largest companies more powerful. Larger companies, like Twitter and Facebook, have the financial means to afford legal teams; as Crowell puts it, "it would really create a mess" (Tempera, 2020).

Section 230, albeit not perfect, is essential to allowing platforms to exist while still within their discretion to do some moderation. Protecting everything from Facebook and YouTube to the recently rebranded 8chan, the central point of section 230 is to provide platforms with the certainty that they can adopt specific, good faith moderation practices that users deem necessary without fear of liability. Without this statute as it stands today, the entire internet environment would look vastly different, some platforms shutting down and others stopping moderation altogether.

4.4 Empirical evidence for a negative societal externality

"We're not just fighting an epidemic; we're fighting an infodemic" - Dr. Tedros Ghebreyesus, WHO Director-General70

Reforming section 230 to account for disinformation draws reasonable skepticism. Notably, Matthew Waxman, a professor at Columbia Law School and specialist on national security law, states that "Disinformation is a broad category that's tough to define, and resulting liability for its harms is very uncertain," suggesting that there is some ambiguity to the negative externalities (Inside Cybersecurity, 2020). Waxman's skepticism is mainly centered around disinformation as it relates to foreign campaigns; however, domestic disinformation "campaigns" pose real, calculable effects on the health and safety of U.S. citizens.

Bursztyn et al. (2020) presents the effects of news coverage of COVID-19 by the two most widely viewed cable news shows in the United States, Hannity and Tucker Carlson Tonight, on viewers' behavior and downstream health outcomes (Bursztyn, 2020). Carlson warned viewers about the severity of COVID-19 from early February, while Hannity originally dismissed the associated risks before incrementally adjusting his position in late February. Via constructed epidemiological models, as exhibited in Appendices C and D, different effect sizes provide empirical evidence that misinformation is an important mechanism driving the observed effects on increased cases and deaths. Given notable, inherent externalities in a pandemic, misinformation may have detrimental effects far beyond those on viewers themselves by affecting disease transmission

5. CONCLUSION

Revise, revoke, repeal – whatever 'R' word you choose – the debate over the future of section 230 boils down to one, simple idea: the law exists to protect people. How can we, as lawmakers and policy drafters, orient the clauses of section 230 such that the statute exists to protect all people? Looking forward, Congress must: a) enact a statutory exemption in section 230(c) that permits civil liability for any platform's complicity in the spread or harboring of disinformation; and b) unambiguously define and recommend good faith steps all ICSPs can put into practice.

Integral to the main objectives of a section 230 modification is to ensure more reliable and truthful content appears, spreads, and lives on online platforms. It would be wrong to assume this is solely a debate over the future of free speech; thus, pleas to revoke section 230 in its totality are both unsubstantiated and grounded in misaligned priorities. Platforms like Facebook and Twitter, for example, are aggressively taking action to suppress the spread of disinformation; however, they are plagued by the subjectivity of their own policy decisions. Hence, there exists a deep desire for heightened collaboration and information sharing between the tech and social media industries and the U.S. government. In the very near future, the path of least resistance in determining how the internet should be governed is to solve these tensions and ambiguities together as a society, in a way people feel is both reasonable and equitable. To do so, lawmakers must continue to invite input from technology companies, citing their individual moderation approaches, as well as input from the public in order to make progressive strides in the forthcoming year.



SUPPLEMENTARY APPENDIX

Appendix A

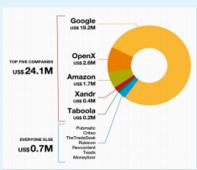
Figure 1: May 26, 2020 tweets made by Donald J. Trump



Notes: Pictured above is the first instance of Twitter flagging tweets made by the President of the United States. Marked with the label "Get the facts about mail-in ballots," users, upon clicking, are taken to a "Twitter Moments" page, which links articles and tweets from journalists and experts that debunk misinformation claims. Brad Parscale, the Trump Campaign manager, labeled this action as a blatant display of Twitter's political bias.75

Appendix B

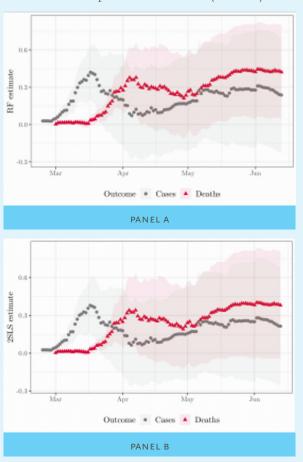
Figure 2: Share of ad revenues generated from COVID-19 disinformation sites



Notes: The top three companies shown (Google, OpenX, and Amazon) generate 95% of the ad revenues to the site in the sample. Google ad services alone delivers \$3 out of every \$4 that these sites earn in ad revenues. Important to note: for Google, the ad shares and revenues for Google AdSense/AdX and Google DFP have been combined.

Appendix C

Figure 3: Reduced-form and 2SLS estimates of effect of differential viewership on cases and deaths (extended)

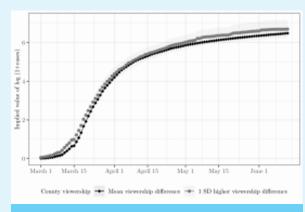


Notes: Figure 3 displays day-by-day reduced form (Panel A) and 2SLS (Panel B) estimates on log one plus cases and log one plus deaths. Panela A portrays day-by-day effects of the instrument, FoxShared, on log deaths and log cases, conditional on state fixed effects and a large set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the country, the log of the country's

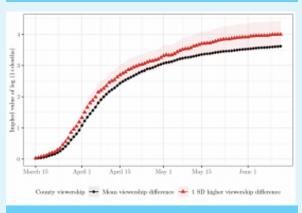
total population, the number of predicted TV's turned to non-Fox channels during Hannity, Tucker Carlson Tonight, and The Ingraham Angle, the population-weighted latitude and longitude, the percent in the country living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. Panel B presents day-by-day effects of the standardized difference in viewership of Hannity vs. Tucker Carlson Tonight, instrumented by FoxShared and controlling for state fixed effects and the same set of covariates as in Panel A. Standard errors are clustered at the DMA level and report 95 percent confidence intervals.

Appendix D

Figure 4: Implied COVID-19 curves



PANELA



PANEL B

Notes: Panel A of Figure 4 plots, in black, the logarithm of (one plus the) mean number of cases in each day across all counties. In gray, the figure plots the implied counterfactual values (based on the 2SLS estimates) for a county with one standard deviation higher viewership difference between Hannity and Tucker Carlson Tonight. Panel B replicates Panel A, taking log one plus deaths as the outcome rather than log one plus cases. This is presented with 95 percent confidence intervals on the counterfactual estimates. Standard errors are clustered at the DMA level.

ENDNOTES

- I. Here, "efficiency" denotes the completion of the core objectives: promoting further development, preserving the free market of ideas, allowing for innovation concurrent with user control, and to do so in a manageable, secure way. But, in a way least reminiscent of George Orwell's "Big Brother" in 1984, yet still prevents the negative externalities of disinformation.
- 2. Important to note, commercial entities should not have the role of determining what constitutes fact over falsehood online. Linking factual articles to tweets gives rise to a multitude of ethical considerations regarding information and perspective biases, beyond the scope of this paper.
- 3. Evaluating platforms like Reddit and Yelp, for example, are key for future study on this issue, yet are outside the scope of this paper.
- 4. Again, ethical considerations arise when the onus falls upon companies, or any governing body, to determine the "right" thing to do. In a way, defaulting to the indivudual companies on platform moderation serves to limit the role and presence of government in our everyday lives.
- 5. For possible future study, it would be of importance to evaluate the various court interpretations of the term "publisher," as it either relates or differs from the term "platform" in both legal provisions and court decisions.
- 6. The nuances behind both qualitative and empirical behavioral theory are crucial to examine in the context of this topic; however, they exist outside the scope of this paper.

REFERENCES

"An Update on Our Political Ads Policy," November 20, 2019. https://blog.google/technology/ads/update-our-political-ads-policy/.

Ardia, David. "Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity under Section 230 of the Communications Decency Act." Loyola L.A. Law Review 43, no. 2 (2010): 373,410.

The New York Times Editorial Board. "Opinion | Joe Biden Says Age Is Just a Number." The New York Times, January 17, 2020, sec. Opinion. https://www.nytimes.com/interactive/2020/01/17/opinion/joe-biden-nytimes interview. html.

Brandom, Russell. "Trump Calls for Last-Minute 230 Repeal as Part of Defense Spending Bill." The Verge, December 2, 2020. https://www.theverge.com/2020/12/2/22037118/trump-section-230-repeal-ndaa-rider facebook-twitter-moderation.

Bursztyn, Leonardo. "Misinformation During a Pandemic." The Becker Friedman Institute of the University of Chicago, June 17, 2020.

Cannon, Robert. "The Legislative History of Senator Exon's Communications Decency Act: Regulating Barbarians on the Information Superhighway." Federal Communications Law Journal 49, no. 1 (November 1996): 52–94.

Cassady, Daniel. "Google Announces Measures to Crack Down On Misinformation Ahead Of Presidential Election." Forbes. Accessed December 16, 2020. https://www.forbes.com/sites/danielcassady/2020/07/31/google-announces-measures to-crack-down-on-misinformationahead-of-presidential-election/.

Twitter. "Civic Integrity Policy," October 2020. https://help.twitter.com/en/rules-and policies/election-integrity-policy.

Communications Decency Act, 47 U.S. § 230 (2018).

Facebook. "Community Standards," 2020. https://www.facebook.com/communitystandards/false_news.

Inside Cybersecurity. "FCC Seeks Input on Regulatory Review of 'Section 230' Liability Protection; Issues Touch on Countering Disinformation," August 4, 2020. https://insidecybersecurity.com/share/11499.

Ghebreyesus, Tedros. "Speech at the Munich Security Conference." Munich, Germany, 2020.

Lotty, Alexandra. "Apps Too: Modifying Interactive Computer Service Provider Immunity under Section 230 of the Communications Decency Act in the Wake of 'Me Too." Southern California Law Review 93, no. 4 (Spring 2020): 855–921.

Mantzarlis. "COVID-19: \$6.5 Million to Help Fight Coronavirus Misinformation." Google, April 2, 2020. https://blog.google/outreach-initiatives/google-news-initiative/covid-19-65-million-help-fight-coronavirus-misinformation/.

Merrill, Jeremy, and Ryan McCarthy. "Trump Won Florida After Running a False Ad Tying Biden to Venezuelan Socialists." Electionland, November 12, 2020.

Global Disinformation Index. "Research - GDI." Accessed December 16, 2020. https://disinformationindex.org/research/.

Sloss, David. "Section 230 and the Duty to Prevent Mass Atrocities." Case Western Reserve Journal of International Law 52, no. 1/2 (Spring 2020): 199-212.

Stratton Oakmont, Inc v. Prodigy Services Co. (Supreme Court, Nassau County, New York 1995).

Tempera, Jacqueline. "Behind the Scenes." Boston College Magazine, Fall 2020.

Thaler, Richard, and Cass Sunstein. Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press, 2008.

Thompson, Alex. "Trump Deploys YouTube as His Secret Weapon in 2020." Politico, September 6, 2020. https:// www.politico.com/news/2020/09/06/trumpyoutube election-comeback-408576.

Volz, Georgia Wells, Robert McMillan and Dustin. "Senate Faults Google, Other Tech Giants for Role in Russian Election Meddling." Wall Street Journal, October 8, 2019, sec. https://www.wsj.com/articles/senate-report-Politics. seeks-to-curb-possible russian-influence-in-2020-election-11570556521.

Working to Stop Misinformation and False News | Facebook Media. "Working to Stop Misinformation and False News," 2020. https://www.facebook.com/formedia/blog/ working-to-stop-misinformation-and-false news.

YouTube - How YouTube Works. "YouTube Misinformation - How YouTube Works." Accessed December 16, 2020. https://www.youtube.com/howyoutubeworks/ourcommitments/fighting misinformation/.

Zuckerberg, Mark. "Testimony of Mark Zuckerberg Facebook, Inc." Testimony presented at the Hearing Before the United States Senate Committee on the Judiciary, November 17, 2020. https://www.judiciary.senate.gov/imo/ media/doc/Zuckerberg%2oTestimony.pdf.

Pulford, J., Hetzel, M. W., Bryant, M., Siba, P. M., & Mueller, I. (2011). Reported reasons for not using a mosquito net when one is available: a review of the published literature. Malaria journal, 10(1), 1-10.

Gawande, A. [@Atul_Gawande]. (2020 November 16) This is the cost of disinformation. It's not only about the reality of deaths but also health & economic damage. MI stats on COVID hospital patients: Typical survivor age = 50-70. At 60 days, 40% either lost their job or were still too sick to work. http://acpjournals.org/doi/10.7326/M2... [Tweet]. Twitter. https://twitter.com/atul_gawande/status/132840 9592959823875?lang=en

Orth, J. [@JodiOrth]. (2020 November 14) I have a night off from the hospital. As I'm on my couch with my dog I can't help but think of the Covid patients the last few days. The ones that stick out are those who still don't believe the virus is real. The ones who scream at you for a magic medicine and that Joe Biden is [Tweet]. Twitter. https://twitter. com/JodiOrth/status/1327771329555292162

Halon, Y. (2020, May 28). Zuckerberg knocks twitter for fact-checking Trump, says private companies shouldn't be 'The arbiter of truth'. Fox News. Retrieved 2020, from https://www.foxnews.com/media/facebook-mark-zuckerberg-twitter-fact-checking-trump

Donald J. Trump. (2020, October 20). Salga a Votar [Vid-YouTube. https://www.youtube.com/ watch?v=BtA1ioujGrQ

About. First Draft. (2022, January 12). Retrieved 2020, from https://firstdraftnews.org/about/

Romano, A. (2018, April 13). A new law intended to curb sex trafficking threatens the future of the internet as we know it. Vox. Retrieved 2020, from https://www.vox.com/ culture/2018/4/13/17172762/fosta-sesta-backpage-230-internet-freedom

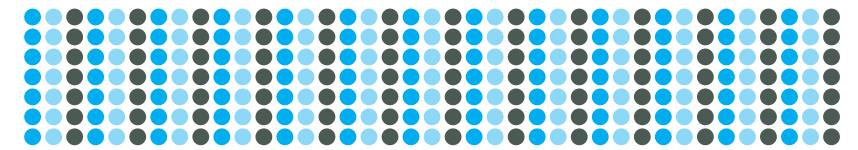
Ingram, M. (2019, August 8). The myth of social media anti-conservative bias refuses to die. Columbia Journalism Review. Retrieved 2020, from https://www.cjr.org/the_ media_today/platform-bias.php

Ellis-Peterson, H. (2018, August 27). Facebook removes accounts associated with Myanmar military. The Guardian. Retrieved 2020, from https://www.theguardian.com/



technology/2018/aug/27/facebook-removes-accounts-myanmar-military-un-report-genocide-rohingya

Stecklow, S. (2018, August 15). Special reports from Reuters journalists around the world. Reuters. Retrieved 2020, from https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/ 25



LIST OF ARTWORK

- T THE BATTLE OF THE SOMME COMIC: THE STORY OF WILLIAM MCFADZEAN
 - © Creative Centenaries. (n.d.). The Battle Of The Somme comic: The story of William McFadzean | CCEA. CEA. Retrieved January 19, 2022, from https://ccea.org.uk/learning-resources/understanding-1916/creative-centenaries/battle-somme-comic-story-william
- I = EASTER PROCLAMATION OF 1916

© originally uploaded to the English Wikipedia by w:User:Jtdirl (https://commons.wikimedia.org/wiki/File:Easter_Proclamation_of_1916.png), "Easter Proclamation of 1916", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-old

28 FIGURE 1

© originally tweeted by User:@realDonaldTrump (account suspended) on 26 May, 2020.

28 FIGURE 2

© GDI report: Ad-funded COVID-19 Disinformation: Money, Brands, and Tech (https://disinformationindex.org/wp-content/uploads/2020/07/GDI_Ad-funded-COVID-19-Disinformation-1.pdf), "Share of ad revenues generated by COVID Disinfo Sites".

FIGURE 3, PANEL A

© originally published in Bursztyn, Rao, Roth, Yanagaziwa-Drott: Misinformation During a Pandemic (https://www.nber.org/papers/w27417), "RF estimates of effect of differential viewership on cases and deaths (extended)".

28 FIGURE 3, PANEL B

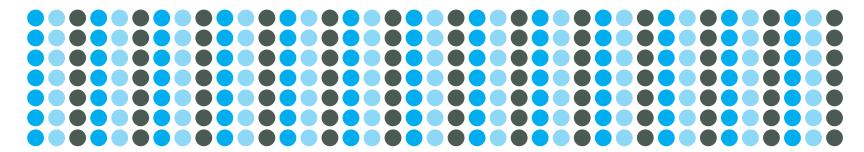
© originally published in Bursztyn, Rao, Roth, Yanagaziwa-Drott: Misinformation During a Pandemic (https://www.nber.org/papers/w27417), "2SLS estimates of effect of differential viewership on cases and deaths (extended)".

FIGURE 4, PANEL A

© originally published in Bursztyn, Rao, Roth, Yanagaziwa-Drott: Misinformation During a Pandemic (https://www.nber.org/papers/w27417), "Implied COVID-19 Curves, Panel A: Estimates on Cases".

2 O FIGURE 4, PANEL B

© originally published in Bursztyn, Rao, Roth, Yanagaziwa-Drott: Misinformation During a Pandemic (https://www.nber.org/papers/w27417), "Implied COVID-19 Curves, Panel B: Estimates on Deaths".

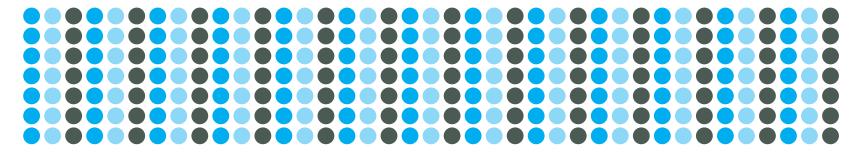


- ARTIFICIAL SYNAPSES BASED ON FERRO-ELECTRIC TUNNEL JUNCTIONS

 © Mike Davies Intel Labs, Intel Corporation et. al: (https://commons.wikimedia.org/wiki/File:Artificial_synapses_based_on_FTJs.png), https://creativecommons.org/licenses/by-sa/4.o/legalcode
- TABLE III INFERENCE POWER AND POWER UTILIZATION

 © Tasbolat Taunyazov, Weicong Sng, Hian Hian See, Brian Lim, Jethro Kuan, Abdul Fatir Ansari, Benjamin C.K. Tee, and Harold Soh: (http://www.roboticsproceedings.org/rss16/po20.pdf), "Table III: Inference Speed and Power Utilization"
- BLACHE LAMONT AT HECLA, MONTANA, WITH HER STUDENTS, 1893
 © Unknown authorUnknown author (https://commons.wikimedia.org/wiki/File:Blanche_Lamont_with_students.jpg), "Blanche Lamont with students", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-US
- THEODORE DURRANT'S MUGSHOT

 © originally published in Report of the trial of William Henry Theodore Durrant (https://archive.org/details/reportoftrialofwoodurr/page/10/mode/2up), marked as public domain.
- © originally published in 31.11 Efficiency and Deadweight Loss (https://saylordotorg.github.io/text_economics-theory-through-applications/s35-11-efficiency-and-deadweight-loss.html), "Figure 31.8 Deadweight Loss"
- GRAPHS SHOWING THE NUMBER AND VALUE OF MERGERS AND ACQUISITIONS
 © originally published in Statista: Mergers and acquisitions in the United States (https://www.statista.com/study/65232/mergers-and-acquisitions-in-the-united-states).
- GRAPHS SHOWING GLOBAL QUERY SHARE BY SEELCT SEARCH ENGINES
 © originally published in Statista: Global market share of search engines 2010-2022 (https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines).
- GRAPHS SHOWING DOMESTIC US SEARCH QUERY SHARE BY SELECT SEARCH ENGINES
 © originally published in Statista: Market share of search engines in the United States 2008-2022 (https://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states).



LIST OF ARTWORK

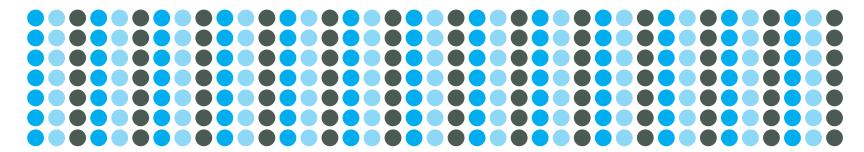
- CORPORATE TECHNOLOGY INFORMATION SERVICES
 - © unpublished tabulation from Corptech Database
- THE BREAKDOWN OF OPINION REGARDING BREAKING UP 'BIG TECH'

 © Statista: Majority of Americans in Favor of Breaking up Big Tech (https://www.statista.com/chart/19440/survey-responses-breaking-up-big-tech)
- FIGURE 1
 © Samadi, Moradi, Moradinazar, Mostafai, Pasdar: Dietary pattern in relation to the risk of Alzheimer's disease: a systematic review (https://link.springer.com/article/10.1007/s10072-019-03976-3), "Figure 2: Healthy and unhealthy diet in relation to AD."
- POTENTIAL SIGNATORIES, SIGNATORIES, AND PARTIES TO THE AIIB

 © L.tak (https://commons.wikimedia.org/wiki/File:Asian_Infrastructure_Investment_Bank_map.svg),
 https://creativecommons.org/licenses/by-sa/4.o/legalcode
- CHINA'S ONGOING TERRITORIAL DISPUTES]

 © Office of the Secretary of Defense (https://commons.wikimedia.org/wiki/File:China's_Disputed_Territories. PNG), "China's Disputed Territories", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-US
- US DEPARTMENT OF JUSTICE AMONG OTHERS ANNOUNCED 23 CRIMINAL CHARGES
 © U.S. Department of Justice (https://commons.wikimedia.org/wiki/File:U.S._Commerce_Secretary_Wilbur_
 Ross,_Acting_Attorney_General_Matthew_Whitaker,_Homeland_Security_Secretary_Kirstjen_Nielsen,_FBI_
 Director_Christopher_Wray_Announces_23_Criminal_Charges_Against_China's_Huawei_&_Wanzhou_
 Meng.jpg), "U.S. Commerce Secretary Wilbur Ross, Acting Attorney General Matthew Whitaker, Homeland Security Secretary Kirstjen Nielsen, FBI Director Christopher Wray Announces 23 Criminal Charges Against China's Huawei & Wanzhou Meng", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-US
- AMERICA GLOBAL TRADE DEFICIT

 © Statista: U.S. Trade Deficit in Goods Reaches Record High, (https://www.statista.com/chart/17281/us-trade-balance).



- CHARLES COUGHLIN, LEADER OF THE CHRISTIAN FRONT
 - © Craine, Detroit (https://commons.wikimedia.org/wiki/File:CharlesCouglinCraineDetroitPortrait.jpg), "CharlesCouglinCraineDetroitPortrait", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-US
- GERMAN AMERICAN BUND RALLY, NEW YORK, MADISON SQUARE GAREN, FEB 1939
 - © Department of Defense. Department of the Army. Office of the Chief Signal Officer. (09/18/1947 02/28/1964); ARC Identifier 36068 / Local Identifier 111-OF-2 1943 (https://commons.wikimedia.org/wiki/File:Nazi_Strike_2.jpg), "Nazi Strike 2", marked as public domain, more details on Wikimedia Commons: https://commons.wikimedia.org/wiki/Template:PD-US