

# Free Speech, Platforms, and the Fake News Problem

Marshall W. Van Alstyne

August 5, 2022

Draft 0.9 – This draft is circulated for comments and critique. Please do forward comments to mva@bu.edu.

## Abstract:

How should a platform or a society address the problem of fake news? The spread of misinformation is ancient, complex, yet ubiquitous in media concerning elections, vaccinations, and global climate policy. After examining key attributes of “fake news” and of current solutions, this article presents design tradeoffs for curbing fake news. The challenges are not restricted to truth or to scale alone. Surprisingly, there exist boundary cases when a just society is better served by a mechanism that allows lies to pass, even as there are alternate boundary cases when a just society should put friction on truth. Harm reflects an interplay of lies, decision error, scale, and externalities. Using mechanism design, this article then proposes three tiers of solutions: (1) those that are legal and business model compatible, so firms should adopt them (2) those that are legal but not business model compatible, so firms need compulsion to adopt them, and (3) those that require changes to bad law.

The first set of solutions, grounded in choice architecture, seek to alter information sets available to those affected by misinformation. By enabling transparency into not simply the content and sources but also the distribution and destination, the system provides effective means for counter narratives that are infeasible under current transparency practices and proposals.

The second set of solutions, based in externality economics, considers how to protect free speech while updating Section 230. Revisions have faced two main critiques: one, that holding platforms actionable for false speech would cause them to take down user speech, and two, that ambiguity in individual messages makes judgement of false speech infeasible at scale. Whistleblower testimony before congress emphasized platform amplification of content in pursuit of engagement. A targeted solution, therefore, can separate original speech from amplified speech, generously protecting the former while reverse amplifying the latter. The posting and even discovery of false speech is protected even as amplification is unprotected. The second element uses scale as an advantage. Rather than vet every message, the system takes only statistical samples. The Central Limit Theorem guarantees that establishing the presence of misinformation in amplified speech is feasible to any level of desired accuracy simply by taking larger samples. A doctor testing for cholesterol does not test every drop of blood but only a statistically valid sample.

The third set of solutions imports insights of antitrust jurisprudence into free speech jurisprudence. The paradox of antitrust before 1978 was that legal decisions, intended to protect consumers and free markets, artificially raised prices by protecting inefficient firms from consequences of competition. Free speech rulings vigorously protect speakers on the basis of enabling a free market of ideas. Overzealous protection of those promoting false facts, however, prevents the market from clearing itself by denying falsehoods competition with truths. No government intervention is required. Rather, it simply needs to step aside in such cases as *WASHLITE v Fox News*, where numerous false stories that covid is no worse than flu and that vaccines do not work have been causally implicated in thousands of unnecessary deaths. The free speech paradox is that legal decisions intended to protect citizens and free idea markets can achieve the opposite, artificially raising harms and cluttering the market with false facts.

# Table of Contents<sup>1</sup>

I.	Introduction .....	- 3 -
II.	What is Fake News? .....	- 5 -
A.	Definitions.....	- 5 -
B.	Intervention Prospects.....	- 7 -
III.	Truth, Falsity & The Nature of Harm .....	- 8 -
A.	False Information Need Not Cause Harm.....	- 8 -
B.	True Information <i>Can</i> Cause Harm .....	- 9 -
C.	Problem Statement.....	- 11 -
IV.	The Value of Free Expression & The Value of Information.....	- 12 -
A.	Goals.....	- 13 -
B.	Value of Information.....	- 15 -
V.	<i>Universal Potentiation</i> .....	- 16 -
A.	Definition .....	- 16 -
B.	Audience vs. Speaker Responsibility .....	- 18 -
C.	Benefits.....	- 20 -
D.	Applying Universal Potentiation: Changing Decisions not Marking Categories - 22 -	
1.	US v Alvarez 567 U.S. 709 (2012) .....	- 22 -
2.	NY Times v. Sullivan 376 U.S. 254, 280 (1964) .....	- 24 -
3.	Brandenburg v. Ohio, 395 U.S. 444 (1969).....	- 31 -
VI.	Solutions Derived from Choice Architecture .....	- 34 -
VII.	Solutions Derived from Pigou .....	- 39 -
VIII.	Solutions Derived from Coase.....	- 41 -
IX.	Utilitarian Solutions to Deontological Problems.....	- 46 -
A.	The Paradox of Free Idea Markets .....	- 46 -
B.	The Paradox of Individual Liberty .....	- 50 -
C.	The Paradox of Preventing Government Tyranny .....	- 55 -
X.	Tests & Objections .....	- 59 -
XI.	Conclusions.....	- 63 -

---

<sup>1</sup> Content in blue is new since the prior version.

## I. Introduction

Fake news is a problem. It is a near universal problem.<sup>2</sup> Trolls and propagandists have used it as a weapon in political campaigns, anti-vaccination campaigns, nutrition battles, insurrections, and to sow ethnic conflict. The inventor of the World Wide Web identified fake news as one of the three most dangerous assaults on the Internet.<sup>3</sup> Fake news is not new. Stone frescoes record the victories of "Rameses the Great" on temples from the 13<sup>th</sup> century BC yet more complete records of the treaty between Egyptians and Hittites show his battle was a stalemate.<sup>4</sup>

Fake news is a known problem. Figure 1 shows occurrences of the terms "false news," "fake news," "misinformation," and "disinformation" in literature since 1800. Until recently, misinformation has been the more common term. Unsurprisingly, "misinformation" rose during both World War I and World War II as conflicting powers sought to demoralize each other's troops with concocted stories of false victories.<sup>5</sup> "Disinformation" spiked during the Cold War between the Soviet Union and the United States, then rose again under the administration of Vladimir Putin.<sup>6</sup> "Fake news" was common under Donald Trump. Deceit is common among adversaries and those pushing hidden agendas.

Challenges of modern fake news, however, render it even more pernicious than in the past. One reason is that modern platforms allow secret or insulated public messaging. The propagandist can whisper his case at a scale that is simultaneously vast yet almost invisible to those who, on observing the message, would oppose it with countervailing evidence. During the 2016 U.S. presidential election, members of the Trump campaign could buy ads targeted at individual coal miners in Pennsylvania.<sup>7</sup> The timing and content of such messages could remain hidden from the Clinton campaign in a manner that is largely impossible using traditional broadcast media.<sup>8</sup> A second reason modern platforms exacerbate the problem is inclusivity. Like most technology, social technology cuts two

---

<sup>2</sup> Governments across all seven continents have been affected. The Wikipedia entry on fake news lists accounts of problems in Australia, Austria, Brazil, Canada, Czech Republic, China, Finland, France, Germany, India, Israel, Malaysia, Mexico, Myanmar, Netherlands, Pakistan, Palestine, Philippines, Poland, Singapore, South Africa, South Korea, Spain, Sudan, Sweden, Syria, Taiwan, Ukraine, United Kingdom, United States, and Venezuela. [https://en.wikipedia.org/wiki/Fake\\_news](https://en.wikipedia.org/wiki/Fake_news) Accessed August 18, 2018.

<sup>3</sup> His other two are citizen surveillance and cyber-warfare. *The World Wide Web's inventor warns it's in peril on 28th anniversary* By Jon Swartz, *USA Today*. March 11, 2017.

<sup>4</sup> Weir, William (2009). *History's Greatest Lies*. Beverly, Massachusetts: Fair Winds Press. pp. 28–41.

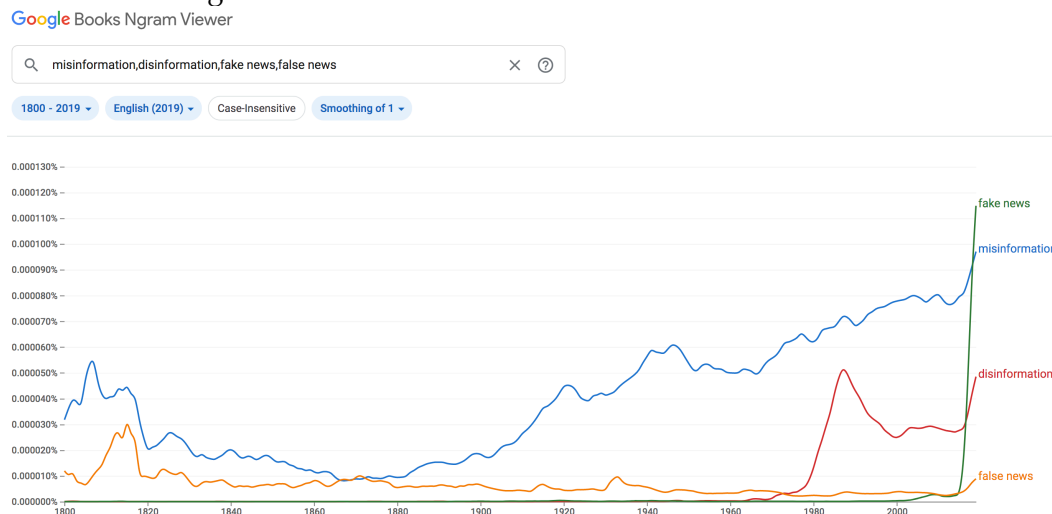
<sup>5</sup> "[Inside America's Shocking WWII Propaganda Machine](#)". December 19, 2016. Retrieved August 19, 2018.

<sup>6</sup> <https://en.wikipedia.org/wiki/Disinformation>.

<sup>7</sup> The Trump campaign introduced disappearing ads on Facebook in Wisconsin, North Carolina and Georgia to deter black votes. Channel 4 News <https://www.youtube.com/watch?v=Klf5ELaOjOk>

<sup>8</sup> The counter argument that Clinton could have bought similar ads is too simplistic. Without knowing what has been said to whom, the cost of covering all possible arguments among all possible listeners is prohibitive. The chief beneficiary, in this case, would be the platform selling blanket advertising.

ways. Modern platforms give voice to the human rights worker, the disenfranchised, the oppressed, the builder, and the whistleblower. They also give voice to the troll, the racist, the enemy state, the bot, and the bot army. Unlike traditional media that hired journalists and broadcasters, modern platforms do not create the content they distribute. Both by law and by design, they absolve themselves of responsibility for propagating fake news, even as they expand the population capable of spreading it. A third reason is false reality born of “deep fake” technology.<sup>9</sup> Victims of political slander could once credibly deny as hearsay false and defamatory claims by third parties. Deep fake technology, however, can create first person fictions that make a victim appear to have committed acts he or she never did or spoken words he or she never would. In appearance and behavior, forged details have become indistinguishable from originals. And, still a fourth reason is the potential for systemic failure. A group of ecologists, biologists, computer scientists, and political scientists warns that social systems, like biological systems, can exhibit cascade failures when pushed too far.<sup>10</sup> Like coral ecosystems that have withstood millennia yet passed in an ecological instant, social systems can break down when pressed beyond their institutional constraints and citizens cannot agree on basic facts.



**Figure 1** – Occurrence of fake news terminology in books and literature 1800-2019. Other terms dominated the term "fake news" until 2016. A rise in misinformation occurred around World Wars I and II, a rise in disinformation occurred during the Cold War, and in fake news in 2016.<sup>11</sup>

<sup>9</sup> Chesney, B., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *Calif. L. Rev.*, 107, 1753.

<sup>10</sup> Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., ... & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27).

<sup>11</sup> Source: <http://books.google.com/ngrams>

Fake news is a hard problem. Open societies tend toward higher growth rates,<sup>12</sup> greater freedom of expression, and greater justice<sup>13</sup> than closed societies. Censorship is a preferred tool of despots and dictators. Freedom of speech is a fundamental right precisely because it exposes those with much to hide. At the same time, a right to free speech must balance a right to privacy and a right to self-rule. Not every fact of one's private life deserves public scrutiny. No lie justifies overturning a fair election. Yet, it can also happen that ugly truths can be more divisive than harmless lies. Design of a fair and balanced news distribution mechanism must weigh rights and properties that conflict with one another yet promote the public good. Presumably, the balance of these rights should be robust to circumstance and not the fragile reactions to politics of the times.

The problem of rights is a problem of balance. Freedom of expression can be used to harass others for exercising their freedom of expression. One right obliterates another. Freedom of expression as a virtue has obvious bounds as a vice.

This article seeks to do three things. The first is to articulate why fake news is a problem and how harm occurs. From this understanding, the second is to lay a foundation for a modern framework of addressing it. Social media platforms of the 21<sup>st</sup> century differ in material ways from the print and broadcast media of the 20<sup>th</sup>. The third is to articulate tests by which one such mechanism might be compared with another. There is no promise that such a framework is either correct or complete, only that it is as balanced as this author can make it.

## II. What is Fake News?

### A. Definitions

A number of scholars have offered a variety of fake news definitions. From their descriptions, a brief summary includes:

- Fabricated information that mimics news media content in form but not in organizational process or intent.<sup>14</sup>
- Fraudulent high velocity content.<sup>15</sup>

---

<sup>12</sup> Przeworski, A., Limongi, F., & Giner, S. (1995). Political regimes and economic growth. In *Democracy and Development* (pp. 3-27). Palgrave Macmillan, London.

<sup>13</sup> Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, 1, 385-472.

<sup>14</sup> Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380) p. 1094.

<sup>15</sup> World Trends in Freedom of Expression and Media Development Global Report 2017/2018. [http://www.unesco.org/ulis/cgi-bin/ulis.pl?catno=261065&set=005B2B7D1D\\_3\\_314&gp=1&lin=1&ll=1](http://www.unesco.org/ulis/cgi-bin/ulis.pl?catno=261065&set=005B2B7D1D_3_314&gp=1&lin=1&ll=1): UNESCO. 2018. p. 202.

- News that is intentionally and verifiably false, and could mislead readers. ... It also includes many articles that originate on satirical websites but could be misunderstood ... when viewed in isolation.<sup>16</sup>
- Presentation of facts that is misleading by design.<sup>17</sup>
- The online publication of intentionally or knowingly false statements of fact ... is a complex and fact-specific endeavor better addressed through case-by-case analysis.<sup>18</sup>
- Disinformation is preferred to 'fake news' and is "...all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit."<sup>19</sup>
- ... is determined by fraudulent content in news format and its velocity.<sup>20</sup> According to Bounegru, Gray, Venturini and Mauri, fake news is when a deliberate lie "is picked up by dozens of other, retransmitted by hundreds of websites, cross-posted over thousands of social media accounts and read by hundreds of thousands" that it then effectively becomes "fake news".
- News from any of three distinct categories: i) stories invented to make money or discredit others, ii) news based on fact but spun to suit an agenda, iii) news that people disagree with.<sup>21</sup>
- Under a narrow view, it is "verifiably false information" but, under a broader view, it includes deliberate attempts at disinformation and news distortion for earning money, promoting ideology or confusion or polarization, or sowing discontent.<sup>22</sup>
- Stories that are provably false, have enormous traction [popular appeal] in the culture, and are consumed by millions of people.<sup>23</sup>
- Fake news is disseminated with the intent to mislead in order to damage an agency, entity, or person, and/or gain financially or politically,<sup>24</sup> often using sensationalist, dishonest, or outright fabricated headlines to increase readership, online sharing, and revenue.<sup>25</sup>

---

<sup>16</sup> Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2), p. 213.

<sup>17</sup> Gelfert, A. (2018). Fake news: A definition. *Informal logic*, 38(1), 84-117.

<sup>18</sup> Klein, D. & Wueller, J. "Fake News: A Legal Perspective" *Journal of Internet Law* 20(10) April 2017. P. 6.

<sup>19</sup> de Cock Buning, Madeleine, A multi-dimensional approach to disinformation : report of the independent High level Group on fake news and online disinformation, Luxembourg : Publications Office of the European Union, 2018. <https://hdl.handle.net/1814/70297>

<sup>20</sup> World Trends in Freedom of Expression and Media Development Global Report 2017/2018 UNESCO p 202.

<sup>21</sup> Newman, N. et. al. Reuters Institute Digital News Report, June 2017.

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3026082](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3026082)

<sup>22</sup> Martens, B., Aguiar, L., Gomez-Herrera, E., & Mueller-Langer, F. (2018). The digital transformation of news media and the rise of disinformation and fake news.

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3164170](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3164170)

<sup>23</sup> Michael Radutzky, producer of CBS *60 Minutes*.

<sup>24</sup> Hunt, Elle. "What is fake news? How to spot it and what you can do to stop it". *The Guardian*; Dec. 17, 2016.

<sup>25</sup> "The Real Story of Fake News" Merriam Webster <https://www.merriam-webster.com/words-at-play/the-real-story-of-fake-news>.

Several scholars, including those of constitutional law,<sup>26</sup> assert that author intent matters.<sup>27</sup> Satire, parody, and entertainment, without deceitful intent, do not constitute fake news according to this view. Interestingly, the 1938 broadcast by Orson Welles of a fake Martian invasion would then not qualify as fake news, despite panic due to belief in its authenticity, as Welles' purpose was entertainment not deceit. Clarifying nomenclature distinguishes "misinformation," which can include "inadvertent sharing of false information" versus "disinformation," which refers to "deliberate creation and sharing of information known to be false".<sup>28</sup>

Tandoc, Zeng & Ling<sup>29</sup> add the role of the audience, further conditioning whether news is fake on the listener's belief as distinct from the author's intent. They argue that without deception, fake news is but a work of fiction. Interrogating subjective and unobservable factors such as author intent and listener belief are challenging but suggest the use of mechanism design to handle information asymmetry. "News you don't believe" is a murky colloquial<sup>30</sup> shortcut. If it is fake news that is disbelieved, then it does no damage and requires no intervention. If it is true news that is disbelieved, then it is not fake but could do damage, an irony that presents design challenges.

Put succinctly, the challenge of designing a fake news intervention based on the forgoing attributes appears to hinge on addressing veracity (falsifiability of news), velocity (speed and reach of news dissemination), and volume (amount and frequency of news production). It may also need to address author intent and listener belief. A leaner and different task, however, will be presented in Section III.C.

## B. Intervention Prospects

Fake news interventions to date have suffered from one or more of four fundamental problems. Technology solutions such as machine learning (ML) and artificial intelligence (AI) suffer from the *arms race* problem. If it is possible to train a filtering system to recognize fake news, it is also possible to train a competing system to recognize when content triggers the filtering system in order to avoid the triggers. The second problem is rating the raters or rather *discrediting the raters*. The actual accuracy of a rater matters little if those who propagate lies succeed in convincing their followers not to believe in the fairness of the

---

<sup>26</sup> Sunstein, C. R. (2021). *Liars: Falsehoods and free speech in an age of deception*. Oxford University Press.

<sup>27</sup> Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. <https://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf>

<sup>28</sup> Wardle, C. (2017). Fake news. It's complicated. *First Draft News*, 16. <https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79>

<sup>29</sup> Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital journalism*, 6(2), 137-153.

<sup>30</sup> Nielsen, R. K., & Graves, L. (2017). "News you don't believe": Audience perspectives on fake news. Reuters Institute for the Study of Journalism. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/ourresearch/news-you-dont-believe-audience-perspectives-fake-news>.

rating system. Liars simply lie about the outcome, seeking to discredit the system that checks their distribution of lies. The third problem is *misplaced responsibility*. A substantial fraction of interventions – crowdsourced fact checking, ML, AI, media literacy, accuracy nudges, tagging, truth chasers, demotion in the news feed – put the burden-of-proof on the platform or on the audience. The typical audience and the platform, however, have less information on content accuracy than the author, implying they are less equipped to police it. By analogy, these solutions place the burden of cleaning pollution costs on those who suffer pollution rather than on those who produce it. The fourth problem is *cost structure*.<sup>31</sup> Making up lies is cheaper than researching hard truths. This leads to a cost advantage for and oversupply of fiction over fact. Deeper solutions will need to address these deeper issues.

### III. Truth, Falsity & The Nature of Harm

#### A. False Information Need Not Cause Harm

Textbooks prior to 2006 published the existence of nine planets. In that year, the International Astronomical Union reclassified Pluto as a dwarf planet,<sup>32</sup> following the 2005 discovery of Eris<sup>33</sup> a celestial body 27% more massive than Pluto. If the larger body was not a planet, then how could the smaller body be one? Texts prior to 1781 suggested there were only six planets.<sup>34</sup> Neither the six nor nine planet claim was true<sup>35</sup> yet the lives of few individuals changed in any meaningful way as new facts corrected old falsehoods.

One Asian fusion restaurant in Cambridge, Massachusetts used the slogan "Eat at Jae's and live forever!"<sup>36</sup> The National Enquirer, a publication with a circulation that reached one million, published a cover story that actress Rita Hayworth had returned from the dead.<sup>37</sup> Each one of these claims is provably false. Few of these claims have altered decisions resulting in consequential damages. None of them, even in total aggregate, have produced social costs rising to a level that would require regulatory oversight.

---

<sup>31</sup> Ibid. Allcott (2017). Ibid. Martens, et. al (2018).

<sup>32</sup> Resolution B5: Definition of a Planet in Our Solar System.

[https://www.iau.org/static/resolutions/Resolution\\_GA26-5-6.pdf](https://www.iau.org/static/resolutions/Resolution_GA26-5-6.pdf) Accessed August 19, 2018.

<sup>33</sup> "Eris is the Greek god of discord and strife, a name which the discoverer Mike Brown found fitting in the light of the academic commotion that followed its discovery." Pluto and the Developing Landscape of Our Solar System. International Astronomical Union <https://www.iau.org/public/themes/pluto/>. Accessed August 19, 2018.

<sup>34</sup> Uranus was discovered March 13, 1781 by William Herschel. Neptune was discovered Sept 23, 1846 by Urbain Le Verrier, Johann Gottfried Galle, and John Couch Adams. Pluto was discovered Feb 18, 1930 by Clyde Tombaugh.

<sup>35</sup> Logically, as long as the definition remains consistent, the number cannot be six or nine. If the definition includes dwarves, then the set of planets includes previously undiscovered bodies Eris and Ceres for a set of at least eleven. If one excludes dwarf planets, then the set includes only eight.

<sup>36</sup> See Boston Globe review posted [http://archive.boston.com/dining/globe\\_review/1087/](http://archive.boston.com/dining/globe_review/1087/). Jae's discontinued using the slogan after several locations closed.

<sup>37</sup> "I'm back from the dead – For two years I was a zombie" *National Enquirer*; Vol 35, No. 14. Dec. 15, 1963



Even the belief in false news need not cause harm.<sup>38</sup> One might believe that vaccines do not work yet still take them in compliance with the law or one might believe the world is flat yet still take cruises that sail the globe. Those beliefs had no ill effect. People have built cathedrals, written songs, sculpted art in the belief that a martyr had risen from the dead.<sup>39</sup> Only when one acts on beliefs that cause harm does fake news lead to a social interest in curbing the harm. Then also, restricting actions can occur without revising beliefs; they represent separate points of intervention with different ethical and efficiency considerations.<sup>40</sup>

In fact, lies can create immense value. Members of the underground railroad, who lied to bounty hunters after passage of the Fugitive Slave Acts of 1793 and 1850, faced fines and imprisonment but helped thousands of refugee slaves escape to freedom.<sup>41</sup> During the Holocaust, villagers of the town Vivarais-Lignon lied to Gestapo officers about the presence of Jewish refugees, risking their own execution, yet they helped desperate people evade Nazi concentration camps.<sup>42</sup> Collectively, such lies saved hundreds if not thousands of people from incarceration, immiseration, and death.

Falsity alone is not a metric that can determine the need for intervention.

## B. True Information *Can* Cause Harm

During World War I, the English language North China Daily News printed allegations that a German factory was rendering human corpses into fats to produce nitroglycerine and lubricants.<sup>43</sup> This story was false and intended to gain allies in the war. During World War II, however, Nazi Propagandist Joseph Goebbels used the truth about these stories to discredit other true stories of German war crimes against the Jews. In effect, he used truthful news to discredit truthful news.<sup>44</sup>

---

<sup>38</sup>Harari, Y. N. (2018). *21 Lessons for the 21st Century*. Random House, ch 17.

<sup>39</sup> Ibid.

<sup>40</sup> What are the intervention considerations for using which? Likely scale of light touch efficiency of limiting action, next level is limiting spread of false beliefs (and limiting action), top level revising acts and false beliefs (and limiting action). False beliefs that cause neither decision error nor externalities do not require intervention. Harm (via either decision error or negative externalities) requires intervention regardless of truth or falsity of news. When does a person need to be protected from self-deception? We don't allow lies about medical products leading to belief in efficacy.

<sup>41</sup> <https://www.history.com/topics/black-history/fugitive-slave-acts>

<sup>42</sup> <https://time.com/5680342/french-village-rescued-jews/>

<sup>43</sup> British Broadcasting Company, "The Corpse Factory and the Birth of Fake News." 17 February 2017. <https://www.bbc.co.uk/news/entertainment-arts-38995205>

<sup>44</sup> Marlin, Randal (2013) **Propaganda and the Ethics of Persuasion**. Broadview Press.

---

*“To tell a truth with ill intent beats all the lies you can invent.” – William Blake  
1919*

---

In 2009, the largest data breach in history took place when a lone hacker broke into Heartland Payment systems and exposed the data of 130 million credit card holders. The practice, common among thieves, is to monetize either directly by billing these accounts or indirectly by selling them to other criminals. The more

accurate and complete the data, the more value it has on the black market and, correspondingly, the greater is the damage to card holders.

In 2018, a US federal court banned the free dissemination of blueprints for 3D printed guns. If the blueprints had been inaccurate or fake, they could not have been used to print working guns. The judge banned them on the basis that harm to the private defendant’s First Amendment rights “are dwarfed” by the harm States might incur if anyone could print a gun.<sup>45</sup> Criminals could defeat security scanners. Printed plastics would facilitate terrorist hijackings. Persons legally barred from gun ownership, due to prior conviction, restraining order, or mental condition could summarily obtain them.

In 2016, Russian troll accounts Blacktivist and DrConservaMom used social media to broadcast true stories of white officers shooting black men and of school shootings. Their messages used true information, tinged with political spin, to suppress black votes in neighborhoods favoring democrats and to animate gun rights voters in neighborhoods favoring republicans. Twitter identified more than 50,000 accounts linked to Russia,<sup>46</sup> and suffered a 21% stock loss after purging more than 1 million fake accounts.<sup>47</sup> Facebook claims have deactivated numerous fake accounts prior to the US 2020 election.<sup>48</sup> It is easy to cull such messages on the premise of foreign interference in sovereign elections. It is not so easy to cull such messages propagated by domestic citizens themselves. Indeed, multiple such accounts were used in the 2016 and 2020 US elections.<sup>49</sup>

Among ideologues and propagandists, it is common practice to take two truths and add them, falsely, to create a lie. A famous celebrity who did die and who did receive a vaccine a few weeks earlier are used to show the danger of vaccination although it had nothing to

---

<sup>45</sup> Judge Robert. S. Lasnik: Washington v. US State Department NO. C18-1115RSL July 31, 2018.

<sup>46</sup> Koh, Y. “Twitter Reveals 1,000 More Accounts Tied to Russian Propaganda Agency,” *Wall Street Journal*. Jan 22, 2018.

<sup>47</sup> Market Watch, “Twitter stock plunges 21% after earnings show effects of fake-account purge” July 27, 2018. <https://www.marketwatch.com/story/twitter-shares-slide-16-after-fake-account-purge-new-rules-in-europe-2018-07-27>

<sup>48</sup> <https://www.cnbc.com/2019/11/13/facebook-removed-3point2-billion-fake-accounts-between-apr-and-sept.html>

<sup>49</sup> Aral, S. (2020). *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health--and How We Must Adapt*. Currency. <https://www.bloomberg.com/news/articles/2020-09-01/social-media-impersonators-seen-as-threat-in-upcoming-elections>

do with his cause of death.<sup>50</sup> Extending their toolkit, they may take three truths and subtract one, again to create a lie. A famous politician who did die of covid despite receiving a prophylactic vaccine is held up as evidence of vaccine inefficacy, omitting the fact that he had cancer of the white blood cells, suppressing his immune system.<sup>51</sup> Antivaxxers routinely cite a small piece of evidence from legitimate research, remove context, and grossly exaggerate it.<sup>52</sup>

Veracity alone is not a metric that can determine the need for intervention.

Much false information causes no harm. Much true information does. Indeed, “truth” is a poor candidate for the operation of law. One cannot own truth. One cannot be liable for truth. One cannot be dispossessed of truth. Laws would have little force if they cannot assign property in truth or liability for truth.<sup>53</sup> Rather, mechanism design should focus on *decisions* for which one can be liable and whose ill-gotten gains may be alienable.

Of forgoing truthful examples that do cause harm, each represents either decision error or a negative *externality*. In the case, of disclosed credit cards, the harm accrues not to the repository or to the thief. Harm accrues to a third party – the card holders whose private data was misappropriated. In the case of Goebbel’s misuse of true information, the harm was not to Goebbels or readers of his fake news but to the Jews who suffered historic atrocities. In the case of working instructions for printed guns, the harm is not to the author of blueprints or the consumer who prints them but to innocent victims who are shot using them. In the case of voter suppression, it is not just the citizens who don’t vote but all other citizens who become governed by a different choice of candidate. In the case of antivaccination exaggerations, not only do the unvaccinated err based on true but incomplete information but also the failure of herd immunity and economic costs accrue to the entire community.

### C. Problem Statement

Beyond the issue of veracity, fake news constitutes a form of information pollution from which there extend potentially large externalities. The need to address externalities as well as consequences of veracity leads to the following definition of the problem.

The phenomenon that this article will address is to clear a communications channel of information that causes harm through decision error or negative externalities when

---

<sup>50</sup> <https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-twitters-ability-to-curb-vaccine-hoaxes> (ibid)

<sup>51</sup>“Vaccine skeptics seize on death of Colin Powell to spread misinformation,” First Draft; Oct 22, 2021. <https://firstdraftnews.org/articles/vaccine-skeptics-seize-on-death-of-colin-powell-to-spread-misinformation/>

<sup>52</sup> Ibid

<sup>53</sup> Calabresi, G., & Melamed, A. D. (1972). Property rules, liability rules, and inalienability: one view of the cathedral. *Harvard Law Review* 85(6), pp. 1089-1128.

propagated at scale.<sup>54</sup> Causes of decision error include false information and true but incomplete information. Causes of negative externalities are involuntary costs imposed on third parties by the voluntary communications activity of any self-interested parties.

This problem statement differs from prior interpretations in two important ways. First, decision error rather than truth per se renders the task easier and separates the problem from intent, which may or may not be discernable. Intent might matter at a penalty phase but is rarely available at the decision to disseminate phase. This problem statement also implicates half-truths that deceive listeners into choosing differently than how they would have chosen under full information. At the same time, it exonerates parody and irony that cause no decision error. Indeed, using irony, the opposite of one's true meaning, is a linguistic device to create emphasis, which can cause the listener to choose more carefully, reducing the chances of error.<sup>55</sup> It also exonerates misstatements that are factually incorrect but not of a magnitude that change a decision. Second, it adds a missing component, harm to third parties. A solution that only envisions harm to the listener fundamentally misses the negative externalities that harm others. Neglecting externalities is one of the profound mistakes made in present day free speech jurisprudence.<sup>56</sup> It is also one of the reasons that platforms, despite seeking to maintain healthy interactions on-platform, have caused devastating harms off-platform.<sup>57</sup> In his testimony before Congress describing Facebook's interaction with Cambridge Analytica, Zuckerberg stated "We did not take a broad enough view of our responsibility."<sup>58</sup> Cambridge Analytica promised games and surveys to Facebook users but in exchange harvested not only all their individual data but also their friends' data, a textbook example of a third-party externality. Friends did not grant permission for data harvesting and were unaware of the breach.<sup>59</sup> The spillover consequence for others is precisely the issue of third-party harm.

#### IV. The Value of Free Expression & The Value of Information

---

<sup>54</sup> This problem statement has a historical parallel in the form of Shannon's channel coding theorem, which seeks to establish whether an information set generated at a source can be communicated without error across a noisy channel to a destination. The answer is yes so long as information volume does not exceed a capacity determined by the noise level. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423. In human terms, the comparable notion to channel capacity is "bounded rationality," a concept central to Section IX.B.

<sup>55</sup> I thank Joel West for this observation.

<sup>56</sup> See The Paradox of Individual Liberty in Section IX.B.

<sup>57</sup> Mozur, P. "A Genocide Incited on Facebook, with Posts from Myanmar's Military" *New York Times*, Oct 15, 2018. <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>

<sup>58</sup> Senate Committee on the Judiciary, Senate Committee on Commerce, Science, and Transportation, "Facebook, Social Media Privacy, and the Use and Abuse of Data." Testimony of April 10, 2018. <https://www.judiciary.senate.gov/meetings/facebook-social-media-privacy-and-the-use-and-abuse-of-data>.

<sup>59</sup> Cadwalladr, C. (2018). I created Steve Bannon's psychological warfare tool': Meet the data war whistleblower. *The Guardian*, 17. <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>

## A. Goals

Judging performance of a fair news platform requires a performance criterion, an optimization metric that has proven elusive in the context of speech. One natural starting point, free speech jurisprudence, has yet to settle on a single goal or metric.<sup>60</sup> Rather, a “deplorable” absence of a unifying goal or theory has left the nature of the Constitution subject to change, often dramatically, as the personnel of the Supreme Court change.<sup>61</sup> Tellingly, adjacent to this 1971 critique, the Supreme Court had just changed the standard for state intervention from “clear and present danger” (*Schenck v United States*, 1919) to “imminent lawless action” (*Brandenburg v Ohio*, 1969), overruling both the tendency to cause sedition or lawlessness (*Whitney v California*, 1927) as well as the test for seditious speech without prospect of imminent action (*Dennis v United States*, 1951). Since that time, the Supreme Court has further moved to shrink state power to regulate false speech (*United States v. Alvarez*, 2012) and to expand corporate power to use money as speech (*Citizens United v Federal Election Commission*, 2010).

The problem of defining a free speech performance metric is hard as it must balance not one goal but many. Enumerated goals include (1) seeking truth (2) supporting free expression (3) participating in government, and (4) stable change in society.<sup>62</sup> Seeking truth, the first goal, confronts the problem that knowledge is subject to continuous improvement as even widely acknowledged truths – a flat earth or geocentric universe – have proven false. Freedom to speak and hear truths is essential to other freedoms. Persons with private or unique information must be permitted to disclose what they know, which promotes growth and spread of knowledge,<sup>63</sup> a tatonnement to truth, and prevention of stagnation and error. Freedom of expression, the second goal, allows people to communicate their needs, wants, and preferences. In juxtaposition with the views of others, free expression offers exploration and affirmation of self, which shapes culture, which in turn shapes self. Participatory governance, the third goal, requires that people have means of protecting their liberties and driving decisions that affect them. “Once one accepts the premise ... that governments derive ‘their just powers from the consent of the governed,’ it follows that the governed must ... have full freedom [to exercise their right of consent].”<sup>64</sup> Participatory expression then offers an important safety valve, supporting the fourth goal, stable social change. Suppressing dissent drives dissenters underground where extra-legal forms of protest become more likely. “Only a government that consistently fails to relieve grievances need fear [political upheaval].”<sup>65</sup> Free expression is a leavening agent fostering political change.

---

<sup>60</sup> More recently, Shiffrin has written that the court’s decisions are “an endless maze” with “no general framework.” See Shiffrin, S. H. (2016). *What's Wrong with the First Amendment*. Cambridge University Press.

<sup>61</sup> Bork, R. H. (1971). Neutral principles and some First Amendment problems. *Ind. Lj*, 47(1), p 1.

<sup>62</sup> Emerson, T. I. (1962). Toward a general theory of the First Amendment. *Yale Lj*, 72, 877.

<sup>63</sup> Balkin, J. (2021) How to Regulate (and Not Regulate) Social Media. *J Free Speech L*, 1(1) p 71-91.

<sup>64</sup> *Ibid.* Emerson p 883.

<sup>65</sup> *Ibid.* p 887.

These four different cases suggest different categorical tests rather than a single organizing principle. Indeed, “effective reasoning by example requires the creation and use of categories through which the lessons of the past can be channeled into service as precedent for the future... Without categories, there can be no rules.”<sup>66</sup> The current use of categories is deontological and not utilitarian.<sup>67</sup> US courts do not balance costs and benefits to decide whether speech is protected, a utilitarian approach. Instead, they examine its case or category to decide a level of protection then, rule to protect it according to that level, a deontological approach. As Chief Justice John Roberts held in *United States v Stevens, 2010*, the “First Amendment’s guarantee of free speech does not extend only to categories of speech that survive an ad hoc balancing of relative social costs and benefits. The First Amendment itself reflects a judgment by the American people that the benefits of its restrictions on the Government outweigh the costs.” Operation of categorical rules that avoid determining person specific value, avoid outcomes unsuitable for general rules.<sup>68</sup> They must operate for all persons equally, as if behind a veil of ignorance, in order that they favor no one specifically. This also has the virtue of blocking government intrusion upon the decision process as such rules are automatic, enjoining state intervention in cases of strictest scrutiny. The imbalance of state power over the individual relative to individual power over the state is such that only by drawing sharp lines around state power to regulate speech can citizens safely balance freedom and authority.<sup>69</sup> Opening the decision process to cost-benefit analysis would only deepen the challenge by adding two further problems: who should decide and on what principled basis? For many, including justices of the Supreme Court, difficulty answering these questions presumptively terminates the utilitarian branch of analysis.<sup>70</sup> Courts eschew valuing free expression, choosing instead to examine content in order to assign category in order to assign protection.

As well justified as these arguments appear, they are nonetheless flawed. The criteria for government intervention since *Schenk* have narrowed and tightened until, as a categorical approach, they have reached a limit. The *Brandenburg* test further enjoined government action for abstractly advocating lawless action. It restricts intervention to speech directed at inciting imminent lawless action *and* a likelihood of producing such action. Despite arriving at an ever tighter boundary, this test exhibits false positives and false negatives. As judged by limiting violence, it proscribes too much.<sup>71</sup> As judged by freeing expression, it permits too little.<sup>72</sup> Expressed in their own terms, paradoxically, these arguments undermine their own goals. A utilitarian approach, grounded in resolving decision errors and externalities, however, addresses these cases.

---

<sup>66</sup> Schauer, F. (1981). Categories and the First Amendment: A Play in Three Acts. *Vand. L. Rev.*, 34, p 265.

<sup>67</sup> Fish, S. E. (2010) “What is the First Amendment For?” *New York Times*, Feb. 1

<sup>68</sup> Fish, S. E. (1999). *The trouble with principle*. Harvard University Press.

<sup>69</sup> *Op. Cit.* Emerson.

<sup>70</sup> *Ibid.* Fish 1999.

<sup>71</sup> See Section V.C.3.

<sup>72</sup> See Section IX.B.

## B. Value of Information

Turning from the value of expression to the value of information proffered by expression, decision theory provides a welcome analytic precision. Information's value is defined as the payoff from an informed decision net of the value of an uninformed decision.

$$\text{Value of Information} = \text{Payoff}(\text{Decision}|\text{Informed}) - \text{Payoff}(\text{Decision}|\text{Uninformed})$$

On learning new information, a recipient updates beliefs and chooses to act or not act according to that choice of action yielding the greatest return. Truth and falsity, goodness and badness, play predictable roles in the outcome. To a recipient, the value of complete and accurate information is always non-negative. It can be zero or greater but never negative. Even bad news allows the recipient to make better informed decisions. If the value is negative, then it must have been incomplete or false.

		Changes Decision?	
		No	Yes
News	Good	<b>A: Value = 0</b>	<b>B: Value &gt; 0</b>
	Bad	<b>C: Value = 0</b>	<b>D: Value &gt; 0</b>

Table 1 – The value of complete and true information is always non-negative. Its value is positive if it changes the default decision.

Information has positive value if and only if it *changes* the default decision. If it represents good news that increases the benefit of a choice one planned to make anyway, then one would have received a windfall benefit whether informed or not. Similarly, if it represents bad news that worsens the outcome of a choice one planned to avoid, then one is no worse off than before. Bad news that a recipient has cancer or that the planet is warming, while unfortunate, nonetheless allow the taking of corrective action to mitigate the damage. Information has value only when it causes the recipient to capture an opportunity that would have been missed, or avoid a loss that would have been captured.

Bad news, if it is accurate and complete, has positive value. By contrast, fake news, because it is inaccurate or incomplete, causes decision error, which produces negative value.

Current jurisprudence focuses on the action or the outcome, separating it from belief on the basis that government either cannot or should not regulate belief but only regulate action. The premise is fair, based on the property that no governance mechanism is infallible in its own beliefs and thus has no basis for imposing its beliefs on others.<sup>73</sup> Intervention is warranted only for actions and then only when one party's actions cause another party's harms.

While the premise is just, its effect on enforcement can be unjust due to a false comparison between the pre *action* and post *action* state as distinct from a comparison between the pre

<sup>73</sup> Op. Cit. Mill; Emerson.

*information* and post *information* state. New information *changes* the pre action and post action payoffs such that the former test condition for harm fails where the latter succeeds.

Consider a portfolio of assets controlled by a fiduciary who receives bad news of a steady and permanent decline. Selling at the earliest opportunity locks in a loss yet prevents a far greater loss. Comparing pre action to post action states shows negative value, the loss, where comparing the decision based on pre informed and post informed states shows positive value, the avoided bigger loss. The informed decision to sell compared to the uninformed decision to hold clearly provides advantage. The opposite condition can also hold. Had the news been good, a steady and permanent increase, then a prompt sale appears to lock in a gain when it misses a far greater opportunity. Comparing states shows positive value where comparing decisions shows negative value. Setting aside externalities, true information *always* produces non-negative value for its recipient. Speech that implicates decisions must therefore focus on the pre and post information states and not the pre and post action states. A judiciary that regulates speech based on comparing pre and post action states is a fiduciary that destroys value.

## V. *Universal Potentiation*<sup>74</sup>

The task is to clear communications of dysfunctional information yet we need a measure of how well the task is done. Doing it badly is the present cause for concern. Thus, the purpose of the present articulation is to identify a goal by which we can measure the success of protecting speech.

### A. Definition

The proposed principle is *Universal Potentiation*, a grant to all persons of such an expansive right of free expression as allows each to reach their full potential, as reflected in their own decisions and those of others, on best available information. This defines an individual welfare criterion. The sharing of information, the action of expression, creates decision options and decision outcomes that determine value. This right of expression is subject only to an equal right manifest in others and responsibility for decision errors that one's free expression induces in others.<sup>75</sup> The "best available information" condition serves to restrict attention to a time and place rather than allow change without limit. These joint and bounded decisions, of speakers and of audiences, thus adhere to the harm principle "that the only purpose for which power can rightfully be exercised over any member of a civilized

---

<sup>74</sup> Potentiation: the action of a substance, at a dose that has no ill effect, that enhances another substance. Potentiate: to make effective or active or more effective or more active. Let expression potentiate decision.

<sup>75</sup> One critique of standard cost-benefit analysis is that a utilitarian agent is too self-focused, which justifies lying too often (see Sunstein, C. R. (2021). *Liars: Falsehoods and free speech in an age of deception*. Oxford University Press). Articulation of *Universal Potentiation* here differs from pure ego centric cost-benefit analysis in one important aspect. If, in extreme cases, one becomes accountable for decision errors induced in others, then social costs are better internalized, so a utilitarian makes a choice closer to the social optimum, rather than one's own personal optimum. The next Section articulates caveats on this accountability.



community, against his will, is to prevent harm to others."<sup>76</sup> To apply *Universal Potentiation* is to ask 'Does expression create options?', 'Does it enable more or better decisions?', 'Do decisions favor the speaker?'

The point of *Universal Potentiation* is first to allow a speaker the chance to move the production possibilities frontier, i.e. to create decision options, and second to allow the chance to claim value from that frontier, i.e. to create decision outcomes. This expanded opportunity set might then serve as the measure of comparing value for free speech jurisprudence in the manner that consumer welfare serves as that comparison for antitrust jurisprudence. Expressive acts that expand the frontier are to be encouraged just as those that contract the frontier are to be discouraged. Likewise, those acts that transfer value voluntarily are to be encouraged while those that transfer value by deceit are to be discouraged.

Critically, the unit of analysis is the set of decisions, if any, attending an expressive act as distinct from the content of the act per se.<sup>77</sup> What decisions does expression potentiate and how do they change? Only when decisions change, either to act or to not act, does speech produce objective differences in value. Identical content can change category, from art to pornography or from copyrighted to public domain or from terrorist propaganda to court evidence, based on context. But if changing context does not change any decision, neither increase nor decrease in value has occurred. By contrast, speech that changes decisions, despite no change in context, will produce measurable value change. A focus on context per se is misplaced.

Likewise, a half-truth that has no effect on a decision neither increments nor decrements its value. A half-truth that changes a decision in the manner of a complete truth has the increase in value of the complete truth. A half-truth that changes a decision in the manner of a complete lie has the decrease in value of the complete lie. The "halfness" is not what matters and a focus on truthfulness per se is misplaced. The decision rather than the content's context or veracity holds the key.

Analysis of expression value must therefore proceed from three questions: What decisions does the expression potentiate? What were the default decisions before the expressive act? How does the expression change payoffs of the informed and uninformed conditions? Knowing these answers yields knowledge of benefits and harms.

---

<sup>76</sup> Mill, John S. (1869) **On Liberty**. Longman, Green, Reader and Dyer publisher.

<sup>77</sup> Ironically, critics of using economic models for legal analysis find fault in that "When we recognize ... decisions ... are not products of logical parthenogenesis born of pre-existing legal principles but are social events with social causes... then we are ready for the serious business of appraising law and legal institutions in terms of some standard of human values." Bush, D. (2000). The Marketplace of Ideas: Is Judge Posner Chasing Don Quixote's Windmills. *Ariz. St. LJ*, 32, p 1147 citing Cohen, F. S. (1935). Transcendental nonsense and the functional approach. *Colum. L. Rev.*, 35, p 847. Thus, our focus on *decisions* here is all the more relevant and consequential.

Several properties become apparent. Decision theory informs us that true and complete information is always value adding to a recipient. Suppression in any manner extinguishes social value. If information of one party is true but incomplete, then dialectic among those with differing views offers far better means of completing the truth than either silence or discourse among those harboring homogeneous beliefs. Veracity is best achieved by diversity. Again, since everyone may be possessed of a missing truth, censoring anyone is a cause of extinguished social value.

That dialectic can correct lies and half-truths is reason to give voice to those of different views. Lies and half-truths are themselves not cause for concern when they cause no decision error. By this principle, irony, parody, and even hateful speech bear not cause for intervention. In fact, irony or parody that cause listeners to become more thoughtful or skeptical reduce their chances of decision error. Linguistically, falsity is used to thwart error not cause it.

Complete and true information always has positive recipient value but it can have negative speaker value. It can give the recipient the option to place burdens upon the speaker. Likewise, new true information can grant power to make demands of the speaker.<sup>78</sup> Increasing options *always* creates recipient value.<sup>79</sup> For the speaker, however, avoiding the associated put burdens and call demands, can motivate half-truths and lies. The mere transfer of value or its creation constitute the benefits of trade. It is the net destruction of value that is the cause of concern.

## B. Audience vs. Speaker Responsibility

If decisions are the proper unit of analysis, who then bears responsibility, speaker or listener for decision errors that follow from speech? Speakers can and will express themselves in ways that favor themselves adverse to others. As a rule, both resource economy and individual agency place primary responsibility for decision consequences upon the decision maker. Listeners are accountable for their own decisions as speakers are accountable for theirs. Personal responsibility thus implies that, between speaker and listener, a listener acting on his or her own behalf should hold him or herself accountable. Principal-agency theory has long established that making a person the “residual claimant” on their decision consequences serves to align incentives, improve effort, and raise welfare.<sup>80</sup> Alignment

---

<sup>78</sup> The former is the equivalent of a ‘put’ option and the latter the equivalent of a ‘call’ option.

<sup>79</sup> Black, Fischer; Scholes, Myron (1973). "The Pricing of Options and Corporate Liabilities". *Journal of Political Economy*. 81 (3): 637–654. Merton, Robert (1973). "Theory of Rational Option Pricing". *Bell Journal of Economics and Management Science*. 4 (1): 141–183.

<sup>80</sup> Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, 3(4), 305-360. Fama, E. F., & Jensen, M. C. (1983). Separation of ownership and control. Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of management review*, 14(1), 57-74. *The journal of law and Economics*, 26(2), 301-325. Grossman, S. J., & Hart, O. D. (1992). An analysis of the principal-agent problem. In *Foundations of insurance economics* (pp. 302-340). Springer, Dordrecht.

avoids the moral hazard that one might indulge in bad decisions when one can shift bad consequences to others. Ensuing incentives also cause people to invest effort to reasonably inform themselves and gather missing information even when speakers are well-intentioned beyond, and in addition to, gathering countervailing information when speakers are ill-intentioned. When needs, wants and wishes are private information, moving the locus of decisions to the informed party can improve decision accuracy. Information revelation usefully occurs when speakers counsel decisions for listeners contrary to those they make for themselves. Such speaker hypocrisy contains its own listener decision value. "What you do speaks so loudly that I cannot hear what you say"<sup>81</sup> would aptly describe, for example, the CEO of an energy firm who advised buying shares in the company for which he was a net seller.<sup>82</sup> A fiduciary role, in which the speaker holds a legal or ethical position of trust relative to the listener, offers useful counterpoint. If the speaker is a fiduciary, then responsibility for decision consequences rests with the speaker. As with doctors, lawyers, and financial planners, a listener must rely on the good faith aid or advice of the fiduciary who is duty bound to avoid conflict of interest and to act in the best interests of the listener. The superior knowledge and expertise of the fiduciary is a principal reason for shifting the locus of decision authority along with culpability to the party in the best position to decide.

Apart from the fiduciary role, a principle of *Universal Potentiation* offers at least five forms of affirmative defense against holding speakers accountable for listeners' decisions. The first is simply that the stated claims are true, or were so, at that time on best available information. Such assertions need not be true at all times and in all ways, only that the speaker had reason to believe they were true when making them. Second, either the truth of the claims or their decision consequences are ambiguous. To protect the freest possible expression, the widest possible latitude must be allowed. Different societies will set the dials on certainty at different points but a default must be a presumption of speaker innocence even when shadowed by the darker shades of doubt. The third is a simple prolepsis. Any advocate can present both sides of an argument when taking a stand. Presenting counterpoint respects the decision process. It alerts listeners to other perspectives, letting them know where to look for more information, and putting responsibility squarely on listeners to make their own decisions. A fourth defense is the absence of undue influence. Freed of pressure, coercion, browbeating, harassment, intimidation, repetition, indoctrination and threats, a listener decides without interference. Only when influence rises to the level of interference may responsibility shift. Finally, a fifth affirmative defense is to show that any misinformation was without consequence. Either no harm occurred or if it did, then misinformation did not alter the default decision. A bad choice that would have been made independent of a speaker's miscues exhibits no information value and thus responsibility for that decision remains with the decision maker.<sup>83</sup> Any single defense is dispositive as it breaks the chain of causal responsibility.

---

<sup>81</sup> Attributed to Ralph Waldo Emerson (cf. Quote Investigator <https://quoteinvestigator.com/2011/01/27/what-you-do-speaks/>)

<sup>82</sup> Hiltzik, M & S. Reckard "Enron Chairman Urged Employees to Buy Stock" LA Times, Jan. 19, 2002. <https://www.latimes.com/archives/la-xpm-2002-jan-19-mn-23640-story.html>

<sup>83</sup> A mnemonic summarizing a test of transfer might be Truth, Ambiguity, Balance, Influence, Change (TABIC).

Taken together, however, the combination of falsity, without alternatives, with undue influence, changing a decision, causing harm, beyond all reasonable doubt is cause for concern.

### C. Benefits

Why would a *Universal Potentiation* standard, as it bears on decisions, be an appropriate standard of care? Three properties suggest it has useful benefits.

The first property is that *Universal Potentiation* is consistent with Mill's articulation of individual liberty.<sup>84</sup> It is utilitarian. It provides for the maximum prospect of self-expression consistent with that same liberty for others and confined only by the principle of harm. The only justification for imposing the will of society on that of the individual is to prevent harm to others. In all but two exceptions, this will provide no *ex ante* restraint to expression instead providing *ex post* remedy. In this, the standard is also content free. No thought or expression is proscribed. Rather it is errors in decisions, adverse to others, that merit scrutiny.

The second property is also utilitarian but from a different perspective. It is consistent with maximum economic value. Each person acting to optimize decisions that affect them resembles the problem of each person acting to optimize consumption that supports them. The First Fundamental Welfare Theorem of economics holds that agents acting in their own self-interest under complete pricing information will compete their way to a Pareto optimum. Everyone will be as well off as possible such that no one can be made better off without making someone else worse off. It requires only knowledge of one's own preferences and posted prices in order to make one's own consumption decisions. It does not require knowledge of others' preferences as these will be revealed by consumption choices that cause prices of scarce resources to rise and those of surplus resources to fall. Each person's decisions contribute to completing the information that becomes an equilibrium.<sup>85</sup> The insight dates to Adam Smith (1776), was expanded by von Hayek (1945),<sup>86</sup> with different versions proven formally by a series of economists dating from Pareto (1906) to Arrow (1951) and Debreu (1951). Exceptions to Pareto efficiency occur when there is information asymmetry in quality, an externality in production or consumption, or monopoly power in price setting. Without irony, these exceptions are precisely the categories indicted by dysfunctional information in the task at hand – clearing a communication channel of information causing decision errors and harms.

The third property is that *Universal Potentiation* is consistent with Emerson's four goals for free speech – truth seeking, self-expression, participatory governance, and stable social change. Either the provision or the seeking of a more complete truth, the first category,

---

<sup>84</sup> Ibid. Mill 1869.

<sup>85</sup> Different initial endowments can yield different equilibria.

<sup>86</sup> Von Hayek, F. (1945) "The Use of Knowledge in Society, *American Economic Review* 35(4), pp 519-530.

necessarily increases value to the recipient. It fosters better decisions and it fosters better science. The articulated principle of *Universal Potentiation* also endorses a maximum of free expression, the second category. It curbs no idea or content at all *ex ante* but rather seeks to limit harms *ex post*.<sup>87</sup> If speech is a tool for achieving other ends – freedom, cooperation, science, governance, education, change, understanding – then rules that blunt this tool risk limiting what we can conceive and what we can accomplish. In terms of governance, the third category, *Universal Potentiation* facilitates influencing *all* decisions that affect one's interests, not just those made by a state. Regardless of whether the audience is a legislature, a judiciary, an executive, a bartender, or a spouse, the principle offers means of providing input to the decisions of others that affect a speaker. Participation in government decisions specifically then provides the safety valve important to the fourth category, stable social change. The rule invites dissent, rather than suppressing it, allowing for redress of grievances and salving the ache that might lead to extra-legal forms of protest. Change can take place via influence upon decisions of individuals and of institutions, in markets and in state rooms, lubricating such progress as negotiation allows. If the purpose of expression is to strengthen not merely current liberty but also future liberty<sup>88</sup> then *Universal Potentiation* can help.

While consistent with these three utilitarian advantages, one inconsistency is that *Universal Potentiation* falls short of a deontological ideal. As expressed by Kant, the *Categorical Imperative* would be to “act only according to that maxim whereby you can at the same time will that it should become a universal law without contradiction.”<sup>89</sup> Extending his analysis, one treats the interests of all other persons as ends in themselves, thwarting one's own self-love and seeking ends that are equal for all persons.<sup>90</sup> Rather, the imperative to “act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end but always at the same time as an end”<sup>91</sup> is a role like that of the fiduciary. That ideal sets a higher bar.

A *Categorical Imperative of Decisions* would have people express such true and complete information as would allow others to decide what is best for themselves. The reduction in information asymmetry, eliminating lemons markets and costly investments in signals and screens, would yield both more and better matches in all forms of exchange. As an ideal, it is an impractical though useful aspiration. It would have each person act as a fiduciary for all other persons even when disclosure is privately costly.

---

<sup>87</sup> Only two exceptions to curbs on *ex ante* expression, as distinct from *ex post* correction will arise from problems of individual liberty and of government tyranny in sections IX.B and IX.C.

<sup>88</sup> Mitchell, G. (2004). Libertarian paternalism is an oxymoron. *Nw. UL Rev.*, 99, 1245.

<sup>89</sup> Kant, I. *Groundwork of Metaphysics of Morals*

<sup>90</sup> Kant, Immanuel (1785). Thomas Kingsmill Abbott, ed. *Fundamental Principles of the Metaphysics of Morals* (10 ed.). Project Gutenberg. p. 23.

<sup>91</sup> *Ibid.* *Groundwork of Metaphysics of Morals*. See Seven Pillars Institute translation <https://sevenpillarsinstitute.org/kantian-duty-based-deontological-ethics/>

*Universal Potentiation* sidesteps the information asymmetry problem that an expressionist does not know, and audience members can misrepresent, the true nature of an audience member's own private interests. Like the First Fundamental Welfare Theorem, which does not require knowledge of others' costs or preferences, a Pareto efficient outcome can be achieved through the decentralized decisions of all. It is not a perfect outcome but it offers a useful standard for what mechanism design can achieve.

#### D. Applying Universal Potentiation: Changing Decisions not Marking Categories

To the extent that measurement is unclear, analysis of three seminal cases illustrates how categorical and utilitarian approaches converge and diverge, raising or reducing social welfare. At all times, the test is to reduce harms caused by communication that induces decision errors or decision externalities. The test of decision change produces fewer false positives and fewer false negatives than tests of intent, of truth, of imminence, and of categories representing "narrowly limited classes of [unprotected] speech".<sup>92</sup>

##### 1. US v Alvarez 567 U.S. 709 (2012)

Xavier Alvarez made lying history on July 23, 2007. At a public hearing of the Walnut Valley Water District Board,<sup>93</sup> Alvarez introduced himself as a retired marine, wounded many times, who had received the Congressional Medal of Honor for bravery. An inveterate liar, he had no such honor.<sup>94</sup> He had never served in the military. A US district court convicted Alvarez for violating the 2005 Stolen Valor Act, which prohibited false claims of having received military honors. A recording of the event left no room for ambiguity. His claims were false. The speaker was knowing.

The outcome of this case was not obvious. The Ninth Circuit appeals court reversed the lower court in a 2-to-1 vote, declaring the Stolen Valor Act unconstitutional. If the government could ban statements merely for falsity, "there would be no constitutional bar to criminalizing lying about one's height, weight, age or financial status on [social media] or falsely representing to one's mother that one does not smoke, drink alcoholic beverages, is a virgin, or has not exceeded the speed limiting while driving on the freeway."<sup>95</sup> The Tenth Circuit increased confusion by finding the Stolen Valor Act constitutional in an unrelated case.<sup>96</sup> The Government appealed the Ninth Circuit decision and lost in a 6-to-3 decision of the Supreme Court.

Different issues weighed in the deliberations of different judges as the Alvarez claims fit no obvious category. Justices of the Supreme Court grappled with the problem that the

---

<sup>92</sup> US v. Stevens, 130 S. Ct. 1577, 1584 (2010).

<sup>93</sup> A branch of the Three Valleys Municipal Water District Board whose records are cited below.

<sup>94</sup> Unrelated lies included having played professional hockey and having once married a starlet from Mexico

<sup>95</sup> Judge Milan Smith, United States v. Alvarez, 617 F.3d 1198, 1200 (9th Cir. 2010), aff'd, 567 U.S. 709, 132 S. Ct. 2537, 183 L. Ed. 2d 574 (2012).

<sup>96</sup> US v. Strandlof, 667 F.3<sup>rd</sup> 1146 (2012).

government could identify neither a specific injury nor an injured party. Absent harm to any person, Alvarez's lie did not qualify as defamation. Not given under oath, it did not qualify as perjury. Dissenting Justice Alito objected that lies proscribed by the Stolen Valor Act legitimately avoided harms such as the award of contracts. False claims had, in fact, defrauded the Department of Veterans Affairs or more than \$1.4 million in stolen veteran's benefits.<sup>97</sup> Alvarez, however, had not sought personal gain and thus his claim could not be marked as fraud. In the end, a court majority found that tests of falsity and of scienter produced false positives, claims that would be punished but should not be punished. Criminalizing false statements of fact would allow the government to create lists of topics that punished inconsequential lying, extending government power without a clear limiting principle. As highlighted in Section III.A, deliberate lies can even raise social value. The villagers of Vivarais-Lignon, who lied to protect Jewish refugees from the holocaust, proved it.

How would one apply a test of harms from decision errors to Alvarez? What decisions changed due to his expression? Three settings prove instructive, one regarding Justice Kennedy's view of Alvarez's introduction, one regarding Justice Sotomayor's interjection of politics, and one regarding Justice Alito's concern for fraud. To proceed, we ask 'what decision is implicated?'. How do defaults change? Do changes cause harm through errors or externalities?

Justice Kennedy held that Alvarez's self-introduction was nothing more than an act of personal expression: "For all the record shows, [his] statements were but a pathetic attempt to gain respect that eluded him. The statements do not seem to have been made to secure employment or financial benefits or admission to privileges reserved for those who had earned the Medal."<sup>98</sup> In this context, his self-introduction did not implicate any specific decision. Having spoken, Alvarez sat down and did not speak again for the remainder of the meeting.<sup>99</sup> His remarks did not bear on any subsequent issue before the water board.<sup>100</sup> Under these circumstances, there is no evidence of a change in *any* decision, not to a vote, not to a payment, not to a contract, nor even to a point of view. Absent a change in decision, the communication was without consequence. Testing for decision harms, intervention is unwarranted. This test result concurs with the Court's decision.

Justice Sotomayor, who voted with the majority, interjected with a contrasting condition during testimony:

*During the Vietnam War, a protester holds up a sign that says, "I won a Purple Heart – for killing babies." Knowing statement: He didn't win the*

---

<sup>97</sup> Justice Alito, dissenting.

<sup>98</sup> Justice Kennedy, writing for the majority.

<sup>99</sup> USA v. Alvarez, Docket No. 2:07-cr-01035 FBI Report Exhibit B.

<sup>100</sup> The public agenda covered commentary on public health goals, approval of minutes, a community relations report, operations report, expense report, legal report, and miscellaneous topics none of which involved Alvarez. See Regular Board Meeting Agenda, July 23, 2007 for the Walnut Valley Water District included in FBI Exhibit B.

*Purple Heart. As a reader, I can't be sure whether he did and is a combat veteran who opposes the war, or whether he's a citizen protesting the war. Is that [citizen protester]... liable under this act?*

The first condition, a veteran who opposes the war, offers a truthful statement. The second condition, a citizen protesting the war, offers an ironic falsehood. Importantly, *both* conditions implicate the same decision: whether to oppose the war. Falsity in the ironic condition does not hide behind deceit but rather draws attention to itself to improve the decision outcome. Neither condition produces harm from decision error and the ironic condition might even prove more persuasive. In the absence of error and externality, intervention is again unwarranted.

By contrast, Justice Alito highlights a third condition implicating fraud. The decision in the context of employment or financial benefit or privilege is whether to grant a benefit based on a lie. If the lie *changes* the default of a benefit denied, then the employer or bank or Veteran's Administration would have a cause of action against Alvarez for fraud. By extension, if Alvarez had displaced an actual Medal recipient by swaying an employer's decision, the externality reflected in the lost job opportunity might also justify a cause of action by the honest veteran. Across these conditions, harm has accrued either in the transfer of an asset to Alvarez or the denial of a benefit to the veteran. The test shows intervention is warranted. Importantly, again, defaults matter. Although Alvarez might have lied in gaining employment, did his lie change the decision? What if he were the only person to apply? If the decision rule is to hire the best person and only one applies then the best is also the worst. Assuming hiring someone is better than hiring no one then the decision defaults to hiring Alvarez. His knowledge of the lie was immaterial. The lie had no consequence at all. The test of decision change exhibits neither false positives nor false negatives across all three sets of deliberations.<sup>101</sup>

## 2. NY Times v. Sullivan 376 U.S. 254, 280 (1964)

The scienter requirement, whether a false claim is knowingly reckless, has become one of the defining tests for whether political speech is actionable. Arising from the Sullivan case, it became "iconic because of its beneficiaries not its reasoning,"<sup>102</sup> suggesting the context and consequences of the "actual malice" standard merit scrutiny compared to alternate tests.

On March 29, 1960, the Committee to Defend Martin Luther King, Jr. took out a full page ad in The New York Times that prompted the Sullivan suit. Dr. King had led a successful boycott of segregated bussing after police arrested Rosa Parks for refusing to give up her seat to a white man. Segregationists in Alabama's state government responded by seeking to silence civil rights advocates through state apparatus. Officials broke up protests, closed

---

<sup>101</sup> This result is also consistent with the verdict in a subsequent case, *Commonwealth v. Crawford* (2021) PA. Super. 102. J-S10033-21, that found a defendant guilty of lying about receiving a Purple Heart in order to join and defraud the American Legion.

<sup>102</sup> McGowan, D. (2022). A Bipartisan Case Against *New York Times v. Sullivan*. *J. FREE SPEECH L.*, 1, p 509.



Alabama State College for Negroes, and filed sham charges. The most recent attempt charged Dr. King with felony tax evasion and perjury for failure to report charitable civil rights contributions on his personal tax returns.<sup>103</sup> The Committee placed the Times ad to solicit funds for Dr. King’s legal bills and to support the civil rights movement.<sup>104</sup> Intended as an emotional appeal, the ad contained several inaccuracies (see Table 2). Times’ editors, relying on the character of those placing the ad, never checked its claims or they would have discovered discrepancies with their own reporting.<sup>105</sup> These errors provided opening for L. B. Sullivan, Commissioner of Police, to sue for defamation even though the ad never mentioned him directly. Decided by an all white jury, under supervision of a judge who had authored segregationist columns for the local paper,<sup>106</sup> the verdict ruled in Sullivan’s favor awarding \$500,000 in damages from the Times. The Alabama Supreme Court affirmed.

**Table 2** – Comparison of false claims in “Heed Their Rising Voices” and historical events.

<b>False Claim</b>	<b>Historical Facts<sup>107</sup></b>
Sang My Country ‘tis of Thee Martin Luther King arrested 7x	Sang National Anthem MLK arrested 4x (charges: loitering, speeding, perjury, and interference with commerce for organizing Montgomery bus boycott <sup>108</sup> )
9 students expelled for leading demonstration at capitol Students protested expulsion by refusing to register	9 students expelled for demanding service at a lunch counter Students protested expulsion by boycotting class
Truckloads of police ringed campus armed with tear gas and shotguns Dining hall padlocked to starve students	Police deployed near campus in large force 3x Unregistered students denied meal tickets

The civil rights issues dwarfed the specifics of the civil suit.<sup>109</sup> Government itself was denying African Americans their civil rights and suppressing press exposure of its efforts. Any remand to the lower courts risked returning equivalent results from the same biased processes that had produced the original verdicts. Any decisions based on technicalities would not clear hundreds of millions of dollars in similar libel lawsuits. The options at hand seemed inadequate.

<sup>103</sup> Lewis, A. (1992). *Make no law: The Sullivan case and the First Amendment*. Vintage Press. See also Edmondson, A. (2019). *In Sullivan's Shadow: The Use and Abuse of Libel Law during the Long Civil Rights Struggle*. University of Massachusetts Press.

<sup>104</sup> Ibid. See also *McKee v Cosby*, Judge Thomas dissenting.

<sup>105</sup> *NY Times v. Sullivan* (1964).

<sup>106</sup> Hall, K. L. (1995). Dignity, Honor, and Civility: *New York Times v. Sullivan*. *OAH Magazine of History*, 9(2), 33-36.

<sup>107</sup> All facts sourced from *NYTimes v. Sullivan* 376 U.S. 254 (1964) unless otherwise noted.

<sup>108</sup> *State of Alabama v. M. L. King, Jr.* No. 7399

<sup>109</sup> Ibid.

What of a test of truth? Factual errors were present, including claims of disgraceful conduct that had not occurred. The common law required truth in all particulars and that standard had not been met.<sup>110</sup> What of a test of negligence? The Times had failed to check the ad against its own reporting, a step that would have caught the principal errors.<sup>111</sup> What of reputation, since any connection “of and concerning” Sullivan was tenuous at best? Dispatching this case did not resolve equivalent cases concerning other state officials who behaved just as badly but were more directly implicated.<sup>112</sup> Governor Patterson, whose libel case was also before the courts, had served on the college board and it was he who had closed the college<sup>113</sup> and initiated the perjury charges when he was state attorney general.<sup>114</sup>

In response, the Supreme Court chose to test intent. Did the press *mean* to publish falsehoods? Was there actual knowledge of or reckless disregard for falsity? Was there “actual malice?” By testing intent, the Court sought to provide “breathing space” for press accounts of official misconduct. Specifically, it held “erroneous statement is inevitable in free debate, and ... must be protected if the freedoms of expression are to have the ‘breathing space,’ [needed] to survive.”<sup>115</sup>

This verdict had no precedent. It ended the use of truth as a defense. It reduced political speech chill but enabled repugnant political discourse. Injurious, false and defamatory claims made with intent to harm but without actual subjective knowledge of falsity or reckless disregard stopped rising to constitutional malice.<sup>116</sup> The Supreme Court ruling in Sullivan “overturned centuries of common law handed down from English courts to extend a right unique to the United States, constitutional protection of speech critical of government, even speech that is false.”<sup>117</sup> As of 2020, the U.S. has adopted a test that no other country has chosen to adopt.<sup>118</sup>

In the process of clearing the defamation suit backlog, the actual malice standard introduced its own problem: Proving subjective knowledge of falsity or recklessness beyond negligence is exceedingly hard, bordering on impossible.<sup>119</sup> Often protecting false claims that do not deserve protection, the test is rife with intervention false negatives. Supreme Court Justice Clarence Thomas describes it as policy masquerading as law. He argues the reasoning is invalid, the history is unprecedented, and the false narratives it allows against public figures

---

<sup>110</sup> Epstein, R. A. (1986). Was New York Times v. Sullivan Wrong. *U. CHI. L. REV.*, 53, 782.

<sup>111</sup> Op. Cit. Lewis.

<sup>112</sup> Ibid.

<sup>113</sup> Hall, Kermit (2004) “L.B. Sullivan” in **100 Americans Making Constitutional History**, M. Urofsky (ed). Pp 189-191.

<sup>114</sup> Op. Cit. Lewis.

<sup>115</sup> Sullivan 376 U.S. 271-272, citing NAACP v. Button 371 U. S. 415, 371 U. S. 433.

<sup>116</sup> Ibid, Kermit 2004.

<sup>117</sup> Edmondson, A. (2019). *In Sullivan's Shadow: The Use and Abuse of Libel Law during the Long Civil Rights Struggle*. University of Massachusetts Press.

<sup>118</sup> Logan, D. A. (2020). Rescuing Our Democracy by Rethinking New York Times Co. v. Sullivan. *Ohio St. LJ*, 81, 759.

<sup>119</sup> Ibid. (Logan)

damage democracy. Lies aimed at democrats, characterizing them as a Satanic child abuse cult, led to a pizza parlor shooting.<sup>120</sup> Supreme Court Justice Gorsuch allows that the verdict might have made sense in 1964 when editors checked facts prior to print but the test no longer makes sense when fact checking plays no role in claims amplified by social media, who are themselves unaccountable.<sup>121</sup> “Those exercising freedom of the press had a responsibility ... to get facts right or, like anyone else, answer in tort for the injuries they caused.”<sup>122</sup> The actual malice standard has become “effective immunity from liability.”<sup>123</sup> Compounding damage from lack of accountability, the actual malice standard reverses journalists’ incentives for integrity. The optimal legal strategy is to publish without investigative fact-checking in order to avoid the scienter requirement.<sup>124</sup> Editors of The New York Times themselves recognized this problem even as they celebrated victory: “we may be opening the way to complete irresponsibility in journalism.”<sup>125</sup> Before she joined the Supreme Court, professor Elana Kagan wrote that the verdict’s effect “may cut against the very values underlying the decision.”<sup>126</sup> If the goal is to inform public discourse in an effort to hold government accountable, flooding the market with misinformation can have the opposite effect.<sup>127</sup> The logic is not even internally consistent: “...to the extent that protecting political speech furthers central First Amendment values ... protecting speech that ... [distorts] voter perceptions ... through dissemination of false information seriously endangers [those values].”<sup>128</sup> Damage, however, does not stop there. If one can impugn the character of public officials without consequence, citizens of lesser character will serve in offices where citizens of better character decline to do so.<sup>129</sup>

From a mechanism design perspective, the effect is the opposite of providing a market of ideas with accurate information. Invariably, in any conflict, the best informed parties are the accuser and the accused. Making accusers unaccountable gives them reason to contrive such accusations as might expand their opportunities at the expense of the accused. They do not internalize costs. They provide more corrupted information. At the same time, the accused are less able to hold accusers to account. The polity loses a credible signal from one who knows the claim is false. Without legal recourse, nothing but hearsay separates the innocent from the guilty. In one stroke, Sullivan created the saint’s burden and the liar’s

---

<sup>120</sup> *McKee v Cosby* No. 17-1256 (1st Cir. 2017), Thomas concurring.

<sup>121</sup> *Berisha v. Lawson*, 141 S. Ct. 2424, 2426, 210 L. Ed. 2d 991 (2021) – Judge Gorsuch Dissenting.

<sup>122</sup> *Ibid.* p 2426.

<sup>123</sup> *Ibid.* p 2428.

<sup>124</sup> *Ibid.*

<sup>125</sup> *Op. Cit. Lewis* (1992). p 219.

<sup>126</sup> Kagan, E. (1993). A Libel Story: Sullivan Then and Now. *Law & Social Inquiry*, 18(1), p 207.

<sup>127</sup> *Op. Cit. Epstein*; *Op. Cit. Logan*; *Op. Cit. McGowan*.

<sup>128</sup> Redish, M. H., & Pereyra, J. (2020). Resolving the First Amendment's Civil War: Political Fraud and the Democratic Goals of Free Expression. *Ariz. L. Rev.*, 62, p 454.

<sup>129</sup> *Op. Cit Hall* (1995).

dividend.<sup>130</sup> The honest politician cannot shed the false accusation. The dishonest politician can hide from the honest accusation.

Defamation involves third parties: the accuser, the accused, and those who would interact with the accused.<sup>131</sup> Absent controls, falsity involves externalities that yield overproduction of damage. The effect of the actual malice standard is both an increase in misinformation and a decrease in tools to clear it. A better solution separates adjudication from remediation. Low cost adjudication makes identification of falsehood simpler and thus the supply of falsehood less valuable. Low cost adjudication also makes signals of falsity credible and thus consumption of falsehood less palatable. Independent of adjudication, courts can adjust remediation for false accusations in accord with a society's desires for warming or chilling speech.<sup>132</sup>

How might low cost adjudication proceed? Put differently, could a different test yield fewer false positives and false negatives? Could a test of *decision-change* perform better than a test of truth, negligence, reputation, or intent?

The Times' ad implicated two decisions: (1) Should a reader contribute to the Martin Luther King, Jr. defense fund specifically and help advance civil rights generally? (2) Should a reader not enter into or maintain an advantageous relationship with Sullivan? In a Universal Potentiation sense, the former expands a speaker's opportunity based on the reader's decision while the latter shrinks another's opportunity based on the reader's decision. The Sullivan trial focuses on decision two. Table 3 presents a more complete set of facts concerning these decisions.

**Table 3 – Historical context at the time of ad placement**

Date	Historical Events
Jan 30, 1956	MLK home firebombed with wife, a fellow parishioner, and baby daughter inside <sup>133</sup>
1956	Attorney general Patterson (later governor) secures a court order barring NAACP activity in Alabama and fining organizations \$100,000 for failing to turn over member names. <sup>134</sup>
Dec 17, 1956	Segregated bussing ruled unconstitutional in <i>Browder v. Gayle</i> , 352 U.S. 903. Organizing the boycott of segregated busses, interfering with commerce, is the charge for which MLK was convicted in <i>State of Alabama v. M. L. King, Jr. No. 7399</i>

<sup>130</sup> Deepfakes create an analogous problem. They provide plausible deniability, allowing the liar to avoid accountability for misdeeds they have done. See Chesney, R., & Citron, D. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Aff.*, 98, 147.

<sup>131</sup> Op. Cit. Epstein.

<sup>132</sup> Op. Cit. McGowan.

<sup>133</sup> Brown, D. L. "After MLK's home was bombed, he refused to back down: This movement will not stop," *Washington Post* Jan, 18, 2021. <https://www.washingtonpost.com/history/2021/01/18/mlk-house-bombed-montgomery-bus-boycott/>

<sup>134</sup> Edmondson, A. (2019). *In Sullivan's Shadow: The Use and Abuse of Libel Law during the Long Civil Rights Struggle*. University of Massachusetts Press.

1959	Sullivan, himself a member of the Ku Klux Klan, defeats an incumbent segregationist to become police commissioner. Expands police civil reserve loaded with Klansmen. <sup>135</sup>
May, 1960	Alabama attorney general Patterson (later governor) charges MLK with felony tax evasion and has him extradited from Georgia. <sup>136</sup>
Feb 28, 1960	Bat-wielding KKK members assault student protestors while state and Montgomery police stand by <sup>137</sup>
Mar 5, 1960	Sullivan issues public statement of intent to use police to “take whatever action might be necessary to disperse [negro troublemakers]” gathering and demonstrating at the state capitol to secure their civil rights <sup>138</sup>
Mar 9, 1960	Sullivan persuades Alabama Governor Patterson to close Alabama State College for Negroes then uses force to break their lunch counter protest <sup>139</sup>
<b>Mar 29, 1960</b>	<b>– Committee to Defend MLK places “Heed Their Rising Voices” ad in NYT –</b>
<b>Apr 19, 1960</b>	<b>– Lester Bruce Sullivan files libel suit in Montgomery AL<sup>140</sup> – – Trial begins Nov 1, 1960 –</b>
May 28, 1960	MLK acquitted of perjury by all white Alabama jury <sup>141</sup>
May 20, 1961	White mob attacks civil rights protesters for 10 minutes with clubs and chains before police arrive. Sullivan had conspired with the mob leader to give time for the siege. <sup>142</sup>

Complete and accurate information concerning Sullivan’s conduct establishes worse conditions than any of the ad’s misstatements indirectly impugning him. He was a confirmed segregationist and member of the KKK, who used members of the KKK as an extension of police authority. He convinced the Alabama state governor to close the Alabama State College for Negroes. His officers stood by while bat-wielding Klansmen assaulted student protestors. By his own hand, he published a document stating intent to suppress free speech, freedom of assembly, and petitioning for redress of grievances by civil rights workers marching on the state capital. He would take “whatever action might be necessary to disperse [negro troublemakers].” The motes of dirt tossed by the ad appear invisible against the sheets of tar hurled by his own behavior.

<sup>135</sup> Hall, Kermit (1995) “Dignity, Honor and Civility: New York Times v. ‘Sullivan’” *OAH Magazine of History*, v9, n2; pp 33-36.

<sup>136</sup> Ibid.

<sup>137</sup> Urofsky, M. “New York Times Co. v. Sullivan,” **Encyclopaedia Britannica**, <https://www.britannica.com/event/New-York-Times-Co-v-Sullivan>. Accessed June 2, 2022.

<sup>138</sup> “[i]f the Negroes persist in flaunting their arrogance and defiance by congregating at the Capitol Sunday the police will have no alternative but to take whatever action might be necessary to disperse them.” Schmidt, C. W. (2014). *New York Times v. Sullivan and the Legal Attack on the Civil Rights Movement*. *Ala. L. Rev.*, 66, p 302. L.B Sullivan, Montgomery Pub. Safety Commissioner Stmt (Mar 5, 1960) <https://digital.archives.alabama.gov/digital/collection/voices/id/3445>.

<sup>139</sup> Op. Cit. Hall 2004.

<sup>140</sup> Ibid. In their capacity as affected government officers, Alabama Governor John Patterson and three others filed similar lawsuits against the same New York Times ad.

<sup>141</sup> *State of Alabama v. M. L. King, Jr.* No. 9595

<sup>142</sup> Op. Cit. Kermit 1995.

More precisely, consider only the information contained in the ad's false claims. Based on full information at the time of Sullivan's suit, would any party who had believed the ad's false claims *change* a decision upon learning just those claims were untrue? Regardless of whether one was a segregationist or a civil rights supporter or a neutral bystander, one would not change a decision to keep an existing relationship with Sullivan or fail to enter a new one based on that alone. The other evidence weighs more heavily. If anything, Sullivan's standing among segregationists improved.<sup>143</sup> Adopting an actual malice standard is unnecessary. The change-of-decision test is unmoved.

Breathing space is not to broaden the boundaries around *intent* to state false claims. Rather, breathing space is consideration of the totality of information in which a specific false claim has no materiality in changing the decision. As with Sotomayor's irony in the Alvarez case, the falsity per se was immaterial.

Subsequent facts add weight to this conclusion. Dr. King was later acquitted of perjury while Mr. Sullivan further conspired with the leader of a white mob to delay police intervention so that the mob might beat civil rights workers with chains. It is worth pausing, however, to note that these events transpired *after* placement of the ad and Sullivan's filing for libel. The change-of-decision test must rest on "best available" information. As this information was not available at the time of filing, it should not be considered no matter how favorable or unfavorable.

Although not raised by Sullivan's lawyers, a fair question may be asked regarding the first decision. Did the ad's false claims rise to a level where charitable giving to the civil rights fund would constitute fraud? More precisely, would a default of not giving change to one of giving, resulting in decision error, based on the false claims alone? As illustration, a claim by a children's cancer charity that 100% of proceeds were spent on hospice care, transport, chemotherapy, and pain medication for children resulted in federal prosecution because less than 3% was spent as promised.<sup>144</sup> There is no evidence that the Committee to Defend Martin Luther King, Jr. spent funds improperly. Nor is there any evidence that any donor provided with full information on civil rights worker beatings, fire bombings, false arrests, and sham prosecutions would have chosen to withdraw their support. The falsity of the claims was immaterial. Relative to best available information, it did not change the outcome in either the primary or secondary decision potentiated by the ad.

The ruling in Sullivan reached the right conclusion for the wrong reason. The malice standard wrongly changes the boundaries around *intervention* when it should rightly change the boundaries around *remediation*. A test of decision change reaches the same conclusion as the Supreme Court in Sullivan but without such a costly increase in false negatives.

---

<sup>143</sup> Justice Black, Op Cit. Kermit 1995 p 190.

<sup>144</sup> Ruiz, Rebecca "4 Cancer Charities are Accused of Fraud," *New York Times*, May 19, 2015.

As a separate issue, the Sullivan case never examined the deeper question of whether The New York Times should be culpable at all. An ad represents a form of paid third party content for which print publishers are liable according to a 1914 law. An Internet platform that had accepted an identical ad would never have been liable according to a 1996 law. The legal discrepancy between falsehoods in print and online publishing is a problem we take up in Solutions Derived from Coase.

### 3. Brandenburg v. Ohio, 395 U.S. 444 (1969)

The purpose in revisiting Brandenburg v. Ohio is not to relitigate falsity – this case never involved lies. Rather the purpose is to highlight the greater efficacy of a test of decision change relative to a test of categories. The test of decision change protects more speech, avoiding false positives, clarifies when intervention is appropriate, and accords better with moral judgment as measured by social harms.

An Ohio court initially convicted Clarence Brandenburg, a leader in a local branch of the KKK, for violating a state statute banning the advocacy of “violence or unlawful methods of terrorism” in pursuit of “industrial or political reform.”<sup>145</sup> Brandenburg’s recorded oratory at an armed Klan rally argued that “if our President, our Congress, our Supreme Court, continues [sic] to suppress the white, Caucasian race, it’s possible that there might have to be some revengeance [sic] taken.” On appeal, the Supreme Court overruled, finding the state statute unconstitutional.

The Supreme Court’s decision in Brandenburg allows prohibition of speech that lies at the intersection of two categories: (1) speech that is “directed at inciting or producing imminent lawless action” *and* (2) speech that is “likely to incite or produce such action.” Brandenburg’s oratory failed the second prong of the test. This dual screen narrowed bounds on intervention from the clear and present danger test on the basis that mere abstract endorsement of illegal activity in pursuit of political ends is protected speech so long as it does not cross the line inducing illegal activity.<sup>146</sup>

Had the KKK engaged in open conflict with members of the African American community, then, an exhortation by Brandenburg “shoot to disable” would pass both halves of the adopted test. It incites imminent lawless violence *and* it is likely to product it. If this were Brandenburg’s actual remark, the test of decision change would support the same conclusion: the decision change from status quo inaction to committing violent action justifies suppressing speech. Yet, unlike the decision change test, the categorical test produces false positives and false negatives, intervening in speech when it should not and not protecting speech when it should.

---

<sup>145</sup> Ohio Rev. Code Ann. §2923.13.

<sup>146</sup> Op. Cit. Brandenburg.

To see this, we transpose the setting to Tulsa, Oklahoma circa June 1921 and reverse the speaker. Seeking to save her family from massacre, an African American woman shouting “shoot to disable” to her husband could reduce atrocities as a white mob murders hundreds of African Americans and burns dozens of city blocks.<sup>147</sup> At one extreme, status quo inaction results in the murder of an innocent woman and her family so that speech driven decision change preserves their lives. At the other extreme, the exhortation “shoot to kill” affects who gets killed without avoiding any killing. A speech directive changing the decision to one of disabling then reduces deaths, even of murderous thugs. The categorical test bars such speech.<sup>148</sup> The decision test protects it.

The decision change test, a utilitarian framing, results in the more moral outcome if by “moral” we preserve more life or avoid more misery.

Though hypothetical, this view is practical. Following Russia’s illegal invasion of Ukraine in February 2022, Facebook changed its policy on posting violent content. In order to help Ukrainian citizens defend themselves, Facebook lifted its proscription on posting methods to commit violence as Ukrainians sought ways to deter Russian aggression.<sup>149</sup> At one level Russian soldiers sought to take control of political institutions but at another level they engaged in mass deportations, sexual violence, torture, and deliberate killing of civilians, leading to indictments for multiple war crimes.<sup>150</sup> Facebook met an immoral invasion with a moral change in policy.

This same decision change test can be used across Alvarez, Sullivan, and Brandenburg and in this final case is more permissive of speech.

A test of decision change exhibits one more essential virtue: it protects viewpoints. Shielding opinions from intervention occurs for the same reason it provides breathing space for trifling falsity. Third party views and immaterial untruths are not “a difference that makes a difference.”<sup>151</sup> They do not change decisions in a manner that also shifts responsibility from listener to speaker. When Brandenburg proclaims, as he did, “Personally, I believe the [African Americans] should be returned to Africa, the Jew returned to Israel,”<sup>152</sup> he has not changed any decision. No one was deported. No one was placed in danger of deportation. The rendering of an opinion cannot shift the burden of a decision from acting-listener to expressive-speaker. Decision makers must take responsibility for their own decisions. They

---

<sup>147</sup> Historical data put these numbers at 100-300 African Americans murdered and 36 city blocks burned. See <https://www.tulsaohistory.org/exhibit/1921-tulsa-race-massacre/>

<sup>148</sup> Alternatively, the categorical test must introduce yet another categorical exception. The point of the current exercise is to avoid yet more epicycles in the law’s view of the solar system of speech and reframe the problem so as to avoid such epicycles entirely.

<sup>149</sup> The change in policy did not permit Ukrainian calls for violence against Russian civilians. See Vengattil, M. and Culliford, E. “Facebook allows war posts urging violence against Russian Invaders” Reuters, March 10, 2022.

<sup>150</sup> Gessen, M. “The Prosecution of Russian War Crimes in Ukraine,” *The New Yorker*, August 2022.

<sup>151</sup> Bateson, G. (1970). Form, substance and difference. *Essential readings in biosemiotics*, 501.

<sup>152</sup> Op. Cit Brandenburg.



cannot transfer that duty to others. That line is not crossed until the speaker's expression involves coercion, until they exercise undue influence in altering the listener's default decision.

"You should kill yourself" said once is an obnoxious but inactionable viewpoint. By contrast, stating it frequently and repeatedly to a mentally ill boyfriend then telling him to get back inside a carbon monoxide filled truck he has just exited establishes causality sufficient to implicate the girlfriend in a suicide when the boyfriend gets back in the truck, *Commonwealth v. Carter* 481 Mass. 352 (2019). These statements pass the undue influence test in transferring decision responsibility from listener to speaker.<sup>153</sup> Based on objective measures of recklessness, would a normal person have recognized the gravity of harm and chosen alternate conduct so as to avoid the harm?

At the same time, a test of decision change easily dispatches one of the chief objections raised by the girlfriend's defense. Her attorneys argued that a guilty verdict would chill speech as doctors and family members might refuse to discuss end-of-life decisions with the terminally ill.<sup>154</sup> But, such an argument presents a false comparison based on the end state, not the change. The proper comparison is between the prospect of a productive life for an eighteen year old, for which termination is value destroying, versus the end of anguish without prospect of recovery, for which termination can be value adding. In the end, what policy is most conducive to liberty? What is the best decision for the individual, by the individual, on best available information? Regardless of one's personal views on any given end state, one must weigh endings against their beginnings to know their value.

Critically, to reemphasize Section V.A, the end condition alone cannot determine value. In both cases above, the same harm occurs but in one it contrasts with an absence of harm and in the other it contrasts with the presence of far more grievous harm. Likewise, the true value of a positive outcome is indeterminate for we must contrast it with a default that might exhibit harm or a far greater Pareto improvement for all. Not decision outcome but decision *change* determines worth.

A final point in favor of testing decision change is that it gives precision to the efficacy of counter speech. If, in a free society, we hold that "the remedy for speech that is false is speech that is true ... the response to the unreasoned is the rational; to the uninformed, the enlightened, to the straight-out lie, the simple truth,"<sup>155</sup> then we must know those occasions when simple truths can defeat straight-out lies. The test of decision change identifies two such occasions: (i) when time remains between the contested expression and the potentiated decision, and (ii) when the decision is reversible. The former has been known for most of the past century and justifies exempting offensive but abstract speech that is protected under both a decision change test and the second half of the Brandenburg test. As Justice Brandeis

---

<sup>153</sup> "... but for the defendant's admonishments, pressure, and instructions, the victim would not have gotten back into [his] truck and poisoned himself to death." *Commonwealth v Carter*. at 635.

<sup>154</sup> *Ibid*, at 636.

<sup>155</sup> *Op. Cit. Alvarez*.

wrote in 1927, "If there be time to expose through discussion the falsehood and fallacies, to avert the evil by processes of education, the remedy to be applied is more speech not enforced silence."<sup>156</sup> The second, however, is less obvious but born of options theory. Decision choices split between reversible and irreversible, defining different mathematical forms for their value.<sup>157</sup> Reversible decisions can be made without perfect information. The option should be exercised quickly. By contrast, irreversible decisions should be made based on best available information. The option should be exercised methodically. For reversible decisions, counter speech can be effective even after the decision has been made, broadening the space where intervention is unwarranted.

Universal Potentiation gives a person the right to increase his or her own welfare via speech to the maximum extent possible subject only to a welfare reduction in others. Analogous to the Consumer Welfare standard whose metric is surplus born of price change,<sup>158</sup> the welfare metric of Universal Potentiation is surplus born of decision change. This has numerous advantages. It protects irony and parody. It gives breathing space to falsehoods that make no difference. It offers viewpoints safe harbor. It defines the boundaries on the efficacy of counter speech. It suffers fewer false positives and false negatives compared to tests of categories.

## VI. Solutions Derived from Choice Architecture

The presence of an externality means that knowledge of the transaction is divorced from knowledge of the harm. The information sets do not overlap. The decision and the action are disjoint. A framework for addressing the problem then admits one of two solutions. The first option is a governance mechanism that moves information about the harm to the party with knowledge of the transaction. Facebook, for example, could learn of damage that its ads cause. The second is a governance mechanism that moves information about the transaction to the party with knowledge of the harm. People affected by ads on Facebook, for example, could learn details of the ads.

The first option is inferior for at least three reasons. (i) Concentrating all information at the center creates a platform of large power and little oversight. At the same time, concentrating off-platform information on-platform has the potential to increase information asymmetry across society creating the potential for widespread exploitation. (ii) Moving information from off-platform to on-platform does not align incentives. Since the party with knowledge of the transaction is *not* the party suffering harm, the central platform need not be motivated to change behaviors so as to improve social welfare. (iii) The near infinite variety of potential

---

<sup>156</sup> *Whitney v California* 274 U.S. 357 (1927), J. Brandeis Concurring.

<sup>157</sup> Dixit, R. K., Dixit, A. K., & Pindyck, R. S. (1994). *Investment Under Uncertainty*. Princeton university press.

<sup>158</sup> Tests used in the US and the European Union, for example, include a Small but Significant and Non-transitory Increase in Price (SSNIP) and Upward Pricing Pressure (UPP). See Farrell, J., & Shapiro, C. (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *Berkeley Electronic Journal*, Vol 10.

externalities ensures that certain forms of damage are likely to be missed. Pulling all possible information with all possible externalities onto the platform is a technical task so daunting that, given any private information among individuals, obtaining complete knowledge is nearly impossible. In effect, the reasons that Facebook “did not take a broad enough view of [its] responsibility” include both that it is technically infeasible and incentive incompatible. Moreover, were it possible to succeed, the outcome would not be desirable.

The second option is superior for reasons that invert the logic above. (i) Moving information that is on-platform to third parties off-platform decentralizes power and reduces information asymmetry. The prospect of exploitation falls. (ii) Parties that suffer harm obtain information on the causes of harm. Incentives align, giving those with the desire to act the information needed to act. Welfare naturally improves. (iii) Public exposure, moving on-platform information off-platform, facilitates parallel search by diverse members of a society. This is decentralized rather than centralized governance. The chances for uncovering the nature of harm improve. At the same time, decentralization fosters a marketplace of information where different ideas compete. Truthful interpretation is easier on the basis of shared facts.

- 1) Access Solution: Pair knowledge of harmful externality with access to means of resolution – grant “equivalence of access” to destinations, not simply transparency into origins, sufficient to allow reversal of decisions.

This is not merely a matter of transparency or maintaining records for a period of years and making the nature of an ad, advertiser, and contact information available. It must also provide access to those who saw the ad and to those where it spread. This was infeasible in print and broadcast media. In terms of *Universal Potentiation*, this solution provides people means of influencing third party decisions in matters that affect them. This provides means to *undo* the damage, not merely allow for its discovery. Information must be actionable not merely knowable. In effect, this rule resembles the “Equal Time Rule” that requires broadcast stations to provide equivalent access to affected parties in order to present their case if requested. *Universal Potentiation* differs, however, in not constraining time allocations at the source. On balance, this solution should be business model compatible as it simply allows one partisan to use the same advertising tools as another partisan.

Resolving information asymmetry between those doing the harm and the locus of harm suggests a first intervention. Media platforms should provide access to those who have received dysfunctional information. That is, media platform should grant an injured party access to media targets equivalent to that obtained by the parties causing injury. Access should include, for example, not only people reached by an ad but also those contacts to whom such ads were shared. Access goes beyond current transparency requirements that record who purchased an ad, its content, and release dates. Transparency only lets affected

---

<sup>159</sup> Astor, Maggie “Now In Your Inbox: Political Misinformation” NY Times Dec. 13, 2021.

parties learn that damage has occurred. It does not provide means of undoing the damage. Limited disclosure only provides information sufficient to reach the perpetrator, not the means to undo harm by enabling counter messaging. By contrast equivalence of access allows an injured party to seek redress not only by holding the perpetrator accountable, which transparency provides, but also by updating facts and narratives among recipients, which access equivalence provides.

Transparency laws that simply record the nature of a communication – its sponsor and its content – do not per se address decision errors or externalities. If a crime were committed, such laws provide only for recording the event or compensating it without undoing the damage. Transparency allows people to discover how they were injured and to hold the speaker accountable. They do not, however, enable the injured party to reach those members of the market for ideas where the false idea has taken root. Transparency might allow one to sue a candidate for a lie placed prior to an election but access can allow a competing candidate to *undo* the lie before that election. If awareness of an externality is insufficient to correct it, then, in order to fix it, access to means must accompany knowledge of ends. Media platforms must therefore record and provide access to recipients of ads received directly as well as those who received them indirectly via sharing and propagation.

Market access need not affect privacy of individual recipients. Neither their identities nor their contact information need be disclosed. Instead, the platform simply mediates access via ads.

Importantly, media platforms should find provision of equivalent market access to be business model compatible. They need not arbitrate truth. They simply sell ads with access to all parties equally. In effect, “Equivalence of Access” on social media resembles the existing FCC “Equal Time Rule” that requires broadcast stations to provide equivalent market access to present their case if requested. “Equivalence of Access” is both distinct from and weaker than the “right of reply” that required free placement of rebuttals on behalf of citizens disparaged in broadcast editorials. The issue in *Red Lion*, a right of reply attached to broadcaster editorializing, hinged on the limited bandwidth. A person who felt attacked by a radio station had few ways to reply and launching a competing radio station was impractical. A legally binding free response might also cause courts to intervene if the broadcaster felt an editorial did not merit a response. Government then needed to adjudicate content. By contrast, this attaches to third parties off-platform. In no case does equivalence of access require the media platform to take a position or speak in a different voice. It only provides a channel for access at published and prevailing rates. As with access to patent pools, terms should be Fair Reasonable And Non-Discriminatory (FRAND). It balances the media platform’s interest in protecting its assets and the social concern with fair markets for ideas. Government plays no role in adjudicating content and, if involved, only decides whether market access is fair.

A right to reply was weakened by the Supreme Court in *Miami Herald Publishing v Tornillo* 418 U.S. 241 (1974) in the case of newspapers. While a law requiring a right of reply did

not prevent editors from saying what they wished, “it exacts a penalty on the basis of the content” and because newspapers are financially limited, “editors may conclude the safe course is to avoid controversy.” By contrast, equivalence of access boosts revenues and, if anything, invites crosstalk among political opponents.<sup>160</sup> Ironically, equivalent access, paid at FRAND rates, has the opposite effect of the court’s concern for an economic burden on the press. Equal access at FRAND rates introduces the moral hazard that a media platform could invite or offer critique in an effort to prompt those affected to purchase a response. It is certainly business model compatible. The truth of each perspective could then surely enter the market.

A partisan might object that a media platform could fill all ad space with lopsided content and thereby protect association with a particular point of view. This might be acceptable for a citizen group that is self-funded for advocacy. It is not acceptable for a group whose function is, in large part, to reach an audience. Such a partisan objection does not apply, as in this case, to platforms funded by advertising. While laudable, the goal of protecting a uniform identity must be subordinated to the goal of providing access to the market of ideas. No society should tolerate discrimination against citizens’ voices on the basis of race or gender. The same holds true for points of view. Transparency resolves information asymmetry regarding knowledge of harm. Equality of access permits the undoing of harms.

One shortcoming of this solution is granularity. It applies at the ad or institutional level rather than the individual level. Transparency of messaging is less feasible and less desirable at the individual connection level. An intervention motivating individual care in sharing higher quality news is thus warranted.

One clear cause of the information pollution problem is that lies spread “farther, faster, deeper and more broadly than truth in all categories of information.”<sup>161</sup> Platforms spread a wildfire of lies in order to build businesses based on engagement. If their goal is to attract eyeballs, then flames will do the trick.

Platforms promote blowhards in the name of newsworthiness. A Cornell study of 38 million articles found that the single greatest source of coronavirus misinformation – including that disinfectant is a cure, that an anti-malaria drug is a cure,<sup>162</sup> that the pandemic was a democratic party hoax, and that masks are not effective at reducing viral spread – was the

---

<sup>160</sup> See Section VIII for an even stronger solution to this problem that reconciles inconsistent liability laws concerning print, broadcast, and internet platforms.

<sup>161</sup> <https://science.sciencemag.org/content/359/6380/1146.full>

<sup>162</sup> Lack of efficacy was never seriously in doubt. Malaria is caused by a unicellular parasite that cannot spread from person to person whereas covid-19 influenza is caused by a highly contagious virus. Cellular organisms and viruses generally differ in their methods of growth, causes of harm, and forms of transmission. Unsurprisingly, antiviral drugs are more effective than antiparasitic drugs in treating the covid-19 virus <https://www.cidrap.umn.edu/news-perspective/2020/10/new-covid-studies-remdesivir-yes-hydroxychloroquine-no>

US president.<sup>163</sup> False claims undermining integrity of the U.S. 2020 election reached the point that they led to insurrection.<sup>164</sup> Reversing this amplification suggests a second solution.

## 2) Reverse Amplification Solution: Add social friction to liars and not just their lies

A straightforward way to implement a “social friction” policy is to selectively reverse amplification. Platforms could adopt a policy that parties convicted of lying will have their social networks trimmed and their messages delayed. Does a person have 100,000 followers? Following a lie, it will be 50,000 and messages will be delayed a week. A badge of dishonor, applied to the liar, can inform followers why the platform no longer pushes that liar’s messages into half the followers’ news feeds. Penalties might apply temporarily for good behavior or increase for bad behavior. Repeated lying could mean having 25,000 followers and messages delayed two weeks. Liars can still say what they wish, even to the point of lying, but then followers would need to go looking for their misinformation in contrast to having the platform promote it.

Social friction has three benefits. First, it directly addresses the problem that lies spread faster, farther and more broadly than truth. It filters pollution at the source. It reduces and delays the channels through which lies spread. It operates at the individual level and does not require monitoring of every message as it imposes a penalty *ex post*. Second, social friction motivates liars to change their behavior. If the goal of a blowhard or ideologue is to attract attention or to move an audience, what better motivator is there than limiting audience access? Labeling and deleting do not work. Unmotivated by truth or integrity, they rephrase and repost. Undeterred and without penalties, ideologues with large networks volley and amplify each other’s false claims.<sup>165</sup> When social friction applies to liars, ideologues choose to shrink their own audiences by telling lies. Networks of echo chambers that willfully propagate lies then self-destruct as they willfully take themselves down. What we have needed is a mechanism that disproportionately weeds out untruths as compared to truths when, up to now, we have had the opposite. Going forward, liars render themselves less potent by limiting their own reach.

The incentive structure highlights a third benefit: This policy places the burden on the proper source, the liar rather than the platform or the reader. Too many attempts at solutions insist that platforms mediate 500 million daily messages<sup>166</sup> that they do not author – do we want them judging every message? Can they? Or, proposals ask readers to sift through mountains of manure to find the truth – will they? Who knows better that a claim is false than the

---

<sup>163</sup> [https://allianceforscience.cornell.edu/wp-content/uploads/2020/10/Evanega-et-al-Coronavirus-misinformation-submitted\\_07\\_23\\_20.pdf](https://allianceforscience.cornell.edu/wp-content/uploads/2020/10/Evanega-et-al-Coronavirus-misinformation-submitted_07_23_20.pdf)

<sup>164</sup> <https://www.usatoday.com/story/opinion/2021/01/13/trump-disinformation-campaign-led-to-capitol-coup-attempt-column/6639309002/>

<sup>165</sup> Zakrzewski, C. “Trump’s Twitter feed is covered in warning labels,” Washington Post. Nov. 5, 2020. <https://www.washingtonpost.com/politics/2020/11/05/technology-202-trump-twitter-feed-is-covered-warning-labels/>

<sup>166</sup> <https://www.internetlivestats.com/twitter-statistics/>

author of that claim? Putting social friction on liars causes authors to think and deliberate before pushing what they know to be false.

Friction does not eliminate false information or ability to speak. Rather, it shifts from all-or-nothing censorship, where information is lost, to a graduated increase in difficulty of dissemination, where information is retained. This is especially useful to society if, in some future condition, the purported falsehood turns out to have been true. The datum is discoverable and the processes used to vet it can be improved.

The economics do not favor platforms voluntarily adding friction. This solution is not intrinsically business model compatible. It is cheaper to produce fake news than true news. By enfranchising everyone, social media platforms shift the balance of supply toward cheaper and therefore more abundant sources of supply. The volume of fake news increases. At the same time, fake news is more engaging. It spreads faster, farther, and more broadly than truth.<sup>167</sup> As demand is higher for novelty, machine learning algorithms push that which generates engagement. Social media embrace this demand. Across production and consumption, a population's news diet shifts. The business model optimizes for profit orthogonal to truth, human health, and institutional health.<sup>168</sup> Senators have chastised such platforms for taking insufficient action against only a dozen individuals that spread up to 65% of vaccine misinformation.<sup>169</sup> Platforms have eschewed reducing politicians' fake news on the basis it is newsworthy.<sup>170</sup> Adding friction reduces engagement which reduces profits.

Addressing the fact that platforms are not well-motivated to self-correct their problems prompts discussion of regulatory solutions for decision errors and externalities. At present, we know of only two solutions for solving externality problems. The first, proposed by Arthur Pigou, levies a tax proportional to the damage in order that the marginal private cost rises to the marginal social cost. The producer then internalizes harms rather than shifting them to the citizenry. The second, proposed by Ronald Coase, creates property rights in the externality and uses markets to trade and price it.<sup>171</sup> This raises the cost of otherwise free disposal, while shifting the burden of harm to whomever can bear it most cheaply. We develop each option in turn.

## VII. Solutions Derived from Pigou

---

<sup>167</sup> Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. (Ibid.) <https://science.sciencemag.org/content/359/6380/1146.full>

<sup>168</sup> Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., ... & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27).

<sup>169</sup> <https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-twitters-ability-to-curb-vaccine-hoaxes>

<sup>170</sup> <https://www.nytimes.com/2019/10/17/business/zuckerberg-facebook-free-speech.html> (ibid)

<sup>171</sup> Hovenkamp, H. (2009). The Coase Theorem and Arthur Cecil Pigou. *Ariz. L. Rev.*, 51, 633.

At least one Nobel Economist has endorsed a variant of the Pigouvian tax intended to focus on the business model rather than the externality per se.<sup>172</sup> Romer suggests applying a Pigouvian tax to digital ad sales. This offers at least three advantages over alternative interventions. First, an ad tax directly alters the business model by favoring ad-free subscription revenue over ad-based third-party revenue. Subscription revenue need not require user tracking. Also, above a minimum participation threshold, subscription revenue does not create an incentive to artificially boost engagement. User privacy could improve even as fake news driven demand declined. Second, a progressive tax, with higher costs for larger firms, could favor entrepreneurial startup. Larger firms created by smaller firm mergers would face larger ad tax bills. Relative to breakup, a progressive ad tax simultaneously solves a market concentration problem normally solved through competition law but does so more effectively. Two half size firms could produce more damage than one full size firm if the principal effect of breakup were to cause each smaller firm to compete more fiercely for user engagement. Not so for the progressive ad tax. Third, an ad tax is not content or consequence based, which avoids all free speech concerns. It removes government from adjudicating speech. And, because it does not involve operational oversight, it reduces risk of regulatory capture that can arise from oversight. Both the Federal Aviation Administration and Interstate Commerce Commission have faced charges of inappropriate ties to the aviation and trucking industries respectively.

### 3) Pigouvian Solution A: Apply a progressive tax to news advertising

Romer's version of a Pigouvian tax has two shortcomings. First, the primary means by which all platforms create value is by consummating matches.<sup>173</sup> They pair people with friends, news, apps, search results, rides, movies, products, and destinations. Effective matching requires tracking. Purported privacy benefits will not fully materialize although subscription revenues do align the interests of user and payer in a way that ad revenues do not. Second, one of its greatest strengths is also its greatest weakness. The damage targeted by an ad tax is unhealthy levels of engagement not fake news per se. By avoiding content issues, the ad tax divorces the levy from the externality. A private subscription service could host antivaxx disinformation, conspiracy theories, and false election narratives but pay no tax, whereas a clean ad-driven service, free of fake news, could pay a heavy tax. For Pigou's solution to work, the penalty must scale with the externality. In the context of factory pollution, a tax on effluent volume encourages the factory to shift to less harmful technology but only if the tax corresponds to the harm. A factory could reduce total effluent yet increase the concentration of pollutant to avoid the tax yet increase the damage. By contrast, if the tax applies to the pollutant specifically, the factory shifts to less polluting technology.

This insight offers a means of reforming a progressive ad tax to improve efficacy: tie the levy to the concentration of harms produced in news effluent. This functions exactly in the

---

<sup>172</sup> Romer, P. "A Tax That Could Fix Big Tech," New York Times; May 7, 2019; Section A, p. 23. <https://www.nytimes.com/2019/05/06/opinion/tax-facebook-google.html>

<sup>173</sup> Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). Platform revolution: How networked markets are transforming the economy and how to make them work for you. WW Norton & Company.



manner of taxing the concentration of harmful effluent in factory production. Yet, a tax on falsehood spillovers acts differently than a tax on specific speech. Testing a statistically valid sample solves three fundamental problems, one of scale, one of accuracy, and one of law. First, one need not certify every message; rather a certification authority need only validate a random sample in order to achieve any confidence level desired. Sampling could even apply to closed chat rooms without violating the privacy of the individuals involved. Second, in a rigorous mathematical sense, a flow rate or aggregation of signals provides a constantly updating Bayesian credibility score. Based on the Central Limit Theorem, larger samples cause estimates of any parameter to converge closer to truth as samples accumulate. This advantage is enormous as it deals even with mixed stories that blend truth with lies. Individual message context becomes far less important because errors only need to fall within an acceptable confidence interval, and this can be determined by sample size. An overarching news credibility score characterizes fitness with respect to the whole message population and not simply a single event. Parties on the left and right might disagree on which messages are true yet agree more readily on the flow rate of truth. Given sufficient statistical samples, consistent deviation from an average score can indicate bias in a specific critic as easily as bias in a specific critique. A third benefit is that the boundaries can scale easily with maturity. A large mature platform might face an effluent rate of .1% while a startup could face one of 1% or 10%. Holding platforms accountable need not disadvantage entrants that lack prior experience. Lastly, a fourth benefit is that a tax on concentration of harm offers a practical means to weaken the liability protections of Section 230 while retaining its broader benefits. The binary choice between either complete liability immunity or accountability for all individual messages is too coarse. Platforms can reasonably object that policing content of 500M individual messages daily is not practical. Societies can reasonably object that policing disinformation that causes political insurrections, deaths from infections, and genocidal riots needs to be practical or the platform should not operate. A reasonable balance – one that can adapt to different societies – is to hold platforms accountable for a specific fraction of effluent. Tax the preponderance of dysfunctional information as distinct from specific instances of disinformation. Measuring a fraction of effluent can exhibit practical scale, adapt to law, and converge to truth.

- 4) Pigouvian Solution B: Tax the platform in proportion to measurable harms. Make it progressive based on size and maturity.
- 5) Section 230 Reform: Separate liability for users' original posts and platforms subsequent amplification, generously protecting the former but reverse amplifying damage from the latter.
- 6) Section 230 Reform: Weaken liability protection for amplification on a statistical rather than individual message basis

## VIII. Solutions Derived from Coase

Truth in markets for political ads is particularly troubling. At one extreme, Facebook defended its decision to take all ads, including outright lies, on the basis that citizens, not

Facebook, should decide.<sup>174</sup> At the other extreme, Twitter chose to ban all political ads to avoid both lies and bias.<sup>175</sup> The former pollutes our discourse with fake news unfiltered. The latter entrenches incumbents and impoverishes discourse. There is a better solution. By combining Coase's ideas on externality markets<sup>176</sup> and information economics signaling, we can create a "market for truth." It requires neither machine algorithms to discern truth nor judgments by a potentially self-interested firm. Instead, it discourages liars from lying. It would work as follows:

- 7) Solution derived from Akerlof, Spence & Stiglitz: Allow ad guarantees in a market for truth. Let anyone wishing to signal the integrity of their claims place a resource at risk as warrant for the truth of those claims.
- 8) Solution derived from Coase: Choose the expected value of the resource at risk to reflect the expected value of social harms.

Give people making strong claims the option (not a requirement) to warrant that their claims are true. Examples include a politician making a claim about an opponent, a policy officer or antivaxxer making a claim about vaccine efficacy, or a consumer product company making a claim about its product efficacy or where it is made (e.g. "made in the USA"). The warrant, posted in advance, serves as a time-limited reward to anyone who can prove the claim is false. To dispute a claim, a challenger pays a modest fee to cover the cost of adjudicating fact-checking. Adjudication is handled by a random sample of peers. Winning challengers claim the warranty to spend as they wish, allowing them to undo the harmful externality. Unchallenged claims, or those judged true, have the warrant returned to the author.

In all cases, the cost of guaranteeing the truth of an honest ad is zero. The false claimant, however, pays for the ad, pays the pledge penalty, and pays in reputation. Simply put, the forfeited pledge is the price of a lie. It is paid only by liars. A politician who wishes to lie may still do so. But lying becomes expensive.

What about the slippery middle ground between truth and falsehood—the innuendo and half-truths that infect so much political advertising? Imagine a photo of Joe Biden and his son looking shifty, accompanied by the tagline: "Hunter Biden served on the board of Ukraine's most corrupt company while his father, as Vice President, did all he could to fire a powerful Ukrainian prosecutor." None of that is exactly false. But it implies the senior Biden tried to prevent the prosecutor from going after the company, when in fact he sought the opposite: he wanted the prosecutor fired for failing to pursue corruption.

---

<sup>174</sup> <https://www.washingtonpost.com/technology/2019/10/17/zuckerberg-standing-voice-free-expression/>

<sup>175</sup> Twitter banned "content that references a candidate, political party, elected or appointed government official, election, referendum, ballot measure, legislation, regulation, directive, or judicial outcome." <https://www.vox.com/recode/2019/11/15/20966908/twitter-political-ad-ban-policies-issue-ads-jack-dorsey>

<sup>176</sup> Coase, R. H. (1960). "The Problem of Social Cost," *Journal of Law and Economics*: Vol. 3 : No. 1 , Article 2.

How should an honest ads market handle an ad like this? It refunds half the pledge for an ad that's half a truth. Based on the egregiousness of the lie, the amount of a refund can correspond to one of the sliding scales fact checkers already use. Indeed, Politifact did rate an ad like the one here as half-true on a scale that ranges from: true, mostly-true, half-true, mostly-false, false, and pants-on-fire. Other fact checkers use similar scales. A market for truth need not be perfect. It just needs to be credible and unbiased. By allowing PACs and politicians to warrant their claims, it changes the balance of power, favoring truth over lies in our political discourse.

In what sense is this a Coasian solution? Why would this work? The extended solution combines externality economics with information economics. A truth market for trading honest ads works for the same reason as a carbon market based on cap and trade. It solves the problem of pricing externalities and markets for trade in externalities already exist. Carbon dioxide is pollution. It is a negative externality that harms others. An entity that is causing damage must pay for that damage by buying pollution credits that put a price on the harm done. Fake news is pollution. It is a negative externality that harms others. The size of the honest ads pledge, that is, the lie price, could be any escrow amount set by the social media platform but should be the expected size of harm done. This negative externality is the "social cost" of the damage done by lying. The crowdsource identification of harm is the market that "trades" the externality. The harmed parties claim the lie price and get paid for the damage they experience. Carbon trading markets work so we can expect markets for truth will also work.

Importantly, a market for truth works even when the amount of damage, the lie price, is not known in advance. Exxon Mobile, for example, might claim that human activity does not cause global warming.<sup>177</sup> The lie price for false ads in the U.S. alone is too small relative to the lie price for global warming internationally. One can quickly see that, if a firm repeatedly pays the lie price, their willingness to keep lying shows private cost is too small relative to social cost. Then, the lie price should rise until they stop the lies causing harm. In other words, we have an "efficient search" process that can force liars to internalize the true social cost of their negative externalities even when that cost is initially unknown. With unknown costs, social efficiency sets the lie price at the expected level of harm<sup>178</sup> yet the starting point can adjust to social taste. Cultures that value freer expression can start the lie price low (even zero) while cultures that value information integrity can start the lie price high. Analogous to car accidents, repeated infractions should raise the liability in either culture. Here, rising liability dissuades liars from repeated lying, finds the efficient price, and limits entrenching the lie.

---

<sup>177</sup> Pierre, J. & S. Neuman "Once again, the U.S. has failed to take sweeping climate action. Here's why" NPR; Oct 27, 2021. <https://www.npr.org/2021/10/27/1047583610/once-again-the-u-s-has-failed-to-take-sweeping-climate-action-heres-why>

<sup>178</sup> Kaplow, L., & Shavell, S. (1995). Property rules versus liability rules: An economic analysis. *Harv. L. Rev.*, 109, 713.

And what about free speech? In the U.S., skeptics might object that an honest ads pledge would not withstand First Amendment scrutiny if the pledge were mandatory. U.S. courts view impediments to speech as violations of free speech. Although this is a uniquely U.S. problem, the system still works even when a pledge is voluntary. If the market for truth is fully functioning, then unwillingness to pledge an honest ad is itself a signal that the author is likely lying because honest ads incur no added cost. The 2001 Nobel Prize in Economics acknowledged the tenets of information economics precisely due to the power of “signals” to separate truth from lies.<sup>179</sup> Informative signals are potentially expensive actions taken by knowledgeable parties that back up their claims. A product sold with a guarantee, for example, is almost always more reliable than a product sold “as is” or “buyer beware.” Good sellers, knowing their claims are true, can offer guarantees that bad sellers, knowing their claims are false, cannot afford to offer. The voluntary signal separates good from bad, and fact from fiction. The proposed mechanism is powerful.

This externality signaling market mechanism exhibits several important properties. (i) Knowing a claim to be false, an author will not want to guarantee its veracity. (ii) True claims are costless to the author, enabling honest authors to voluntarily signal by offering a guarantee. (iii) Together, these properties yield a separating equilibrium based on authors’ private knowledge, distinguishing misinformation from authoritative information. (iv) Initial burden for deciding truth rests with the author rather than the platform or the uninformed reader. This is more socially efficient as it solves a negative externality problem. Specifically, the decision to warrant / not warrant a claim places initial burden of proof on the polluter rather than on third parties compensating for that pollution, thus it should reduce pollution. (v) Crowdsource detection of falsehoods scales. Members of the crowd are motivated to detect false claims by the reward. (vi) The challenge fee discourages false challenges. It also covers the costs of adjudicating a challenge so the mechanism is financially self-sustaining. (vii) Jury adjudication makes it harder for ideologues to discredit fact checking relative to standing bodies whose verdicts, though true, they dislike. (viii) The entire mechanism is decentralized and market based. This solves the conflict-of-interest agency problem of having the platform or an authoritarian government adjudicate truth.

Interestingly, establishment of a “market for truth” is also business model compatible. It removes responsibility for adjudicating truth from the social media platform, returning this to society, yet it enables the platform to participate in the advertising and escrow markets. Conditional on building the institutional infrastructure necessary to support a truth market, an extended Coasian solution combined with information economics is economically sustainable. Such a market, analogous to that for carbon trading, could address the dysfunctional information problem.

---

<sup>179</sup> Spence, M. (1978). Job market signaling. In *Uncertainty in economics* (pp. 281-306). Academic Press. Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press. Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American economic review*, 71(3), 393-410.

To further illustrate their power, truth market warrants can help address the longstanding media liability and speaker reach problems in paid advertising. Inconsistencies in print, broadcast, and internet media liability illustrate legislative advertising policies that emerged based on technologies of their times rather than any deeper principle. Each medium faces different rules. Print publishers face the strictest ad liability.<sup>180</sup> They are free to use their discretion to accept or reject ads but if they print an ad containing false content, they can be held liable.<sup>181</sup> Broadcast media face a “no censorship, no liability” rule for political candidates.<sup>182</sup> They are not free to reject candidates’ ads but in exchange for this lack of freedom, broadcasters also face no liability for candidates’ false claims. Interestingly, broadcaster liability changes for non-candidate advertising. They are free to reject ads from non-candidates but if they accept ads with false claims, then they can be held liable.<sup>183</sup> The third and most privileged category are the Internet platforms who have both discretion over what to accept and immunity from liability regardless of their decisions.<sup>184</sup> These inconsistencies in liability for paid advertising apply independent of the fact that all three categories – print, broadcast, and platform – depend heavily on ads for revenue. Across all three, political interests that are not candidates can be denied access to a given audience if their message, however true, conflicts with editorial policies of the medium. Discourse diversity suffers. Is there means to balance interests of all parties – politics, media, and citizens?

When transaction costs are high, liability for pollution should rest with the lowest cost avoider.<sup>185</sup> Markets can trade their way to efficiency *unless* these costs are high. With low negotiating costs, as between a city and nearby factory spewing foul air, assigning property rights to either party matters little if one can buy rights from the other. Trade regarding moving the factory, moving citizens, or changing the technology will be more efficient than from government fiat, especially if either side has hidden costs or benefits that a government cannot observe. Assigning property rights *does* matter if negotiations must take place between each citizen and the factory. Then assigning liability to the factory, the lower cost avoider, reduces total transaction costs, while still enabling trade. Media that spews misinformation is a lower cost avoider than is each citizen who may or may not know ad veracity and may or may not have resources to check it. Should the medium, which is paid to disseminate an ad, check that ad’s claims one time or should 100,000 citizens, who are not paid, check that ad 100,000 times? Social efficiency demands the former, else it asks 100,000 people to wear gas masks or purify their water rather than ask one factory to clean its effluent.

---

<sup>180</sup> Federal Trade Commission Act of 1914.

<sup>181</sup> Sableman, M. “When politi-fact veers into politi-fiction, broadcast and internet ad rules diverge sharply,” Nov. 27, 2019. <https://www.jdsupra.com/legalnews/when-politi-fact-veers-into-politi-46971/>

<sup>182</sup> Communications Act of 1934.

<sup>183</sup> Ibid, Sableman 2019.

<sup>184</sup> §230, Communications Decency Act of 1996.

<sup>185</sup> Ibid. Coase, 1960. Ibid. Calabresi, 1972. Ibid. Kaplow, 1995.

Between an ad's author and an ad medium – print, broadcast, or internet – who should bear liability? How can a society balance a speaker's interest in reach, a medium's interest in editorial autonomy, and an audience interest in hearing diversity? Truth market warrants offer a simple and principled solution. On the principled basis of Coase, Calabresi and Melamed, let liability vest initially with the medium. But, if an ad's author will warrant a claim, the medium must accept and disseminate the message. In exchange, the author assumes liability absolving the medium. The medium is free to disseminate any message it likes but retains liability for any ad the author does not warrant. Political interests may reach any audience merely by assuming liability for their claims. Media are free to accept or reject any unwarranted ad. Media retain liability for ads the author will not warrant and from which they profit. The party with editorial authority retains liability. The transactions burden is not shifted to citizens, placing solution of the pollution problem onto the lowest cost avoider. Diversity of discourse improves as speakers willing to warrant their claims can reach any audience. Truthfulness rises. Polarization falls.

## IX. Utilitarian Solutions to Deontological Problems

Where courts apply deontological rules to speech without regard to consequence, three absolutist paradoxes arise, one in free markets, one in individual liberty, and one in preventing government tyranny.

### A. The Paradox of Free Idea Markets

A problematic but illustrative case provides realism. The Washington League for Information Transparency and Ethics (WASHLITE) sued Fox News on the basis of propagating false pandemic information on the health effects of precautions such as masking and lockdowns and that COVID-19 warnings were a hoax propagated by political opposition.<sup>186</sup> Despite praising the aims of the case, the court moved to dismiss, citing *US v Alvarez* that falsity alone is not an unprotected category but must be knowing and reckless (*New York Times v Sullivan*). Affirming the dismissal, the court of appeals added that Fox News' statements pertained to policies "clearly [implicating] matters of public concern and [receiving] special First Amendment protections no matter how outrageous," and "... however laudable WASHLITE's intent, it's claim is barred by the First Amendment."

The plaintiff's claim was problematic because it sought principally to bar publication of further misinformation and to issue specific retractions. A further claim, that Fox News violated the Consumer Protection Act, was also found invalid on the basis that cable news did not fall within the ambit of trade, which if applied would still run afoul of First Amendment protections.<sup>187</sup> These two points constitute the currently prevailing analysis. A proposed analysis proceeds by comparing the aims sought in *WASHLITE v Fox News*,

---

<sup>186</sup> *WASHLITE v Fox News* 2020.

<sup>187</sup> *Id.*

enjoining or retracting speech and applying product liability, to those proffered under *Universal Potentiation*, what decisions are implicated and how do defaults change.

Enjoining speech *ex ante* is not a position supported by the principle of *Universal Potentiation* nor is it one supported by the courts. Holding expression accountable *ex post*, however, is where utilitarian policy diverges from deontology. Clear decisions potentiated by Fox News expressions involved whether to take precautions against the spread of a virus during a pandemic. This may involve subordinate, complementary, or interfering decisions such as whether to wear a mask, whether to honor a lockdown, whether to wash hands, or whether to interact in a socially distant manner. A clear set of default decisions was also provided by authoritative sources such as the Center for Disease Control, National Institute of Health, the US Surgeon General, the head of the Whitehouse Coronavirus Task Force Dr. Anthony Fauci, and the World Health Organization, an international body without partisan ties to the state. Against this largely coherent set of defaults, was Fox News persuasive in changing decisions?

Statistical analysis shows a causal link between COVID-19 mortality and viewing the top rated Fox News program,<sup>188</sup> a causal link between viewing Fox News and non-compliance with social distancing rules,<sup>189</sup> and reduced consumption of hand sanitizer and masks.<sup>190</sup> Subsequent to the case, research also documented a causal link to lower vaccination rates.<sup>191</sup> Notably, Fox News showed no historical effect on prior flu vaccination rates, suggesting its effects were COVID-19 specific.<sup>192</sup> Having earlier run stories on reemergence of Polio due to vaccine hesitancy,<sup>193</sup> Fox News had no history of opposing measles vaccines, polio vaccines, or smallpox vaccines. Under a principle of *Universal Potentiation*, Fox News would be free to express itself in whatever manner it preferred to attract viewers just as the classes of persons who had lost family members and institutions who had lost employees would be free to hold Fox News accountable for decision errors and externalities based on their advice. Information subsequent to the initial case also reveals that Fox News statements were knowing. Its hypocritical position opposing a government policy mandating either vaccination or testing among employees stands in contrast to its own

---

<sup>188</sup> Bursztyn, Leonardo and Rao, Aakaash and Roth, Christopher and Yanagizawa-Drott, David, "Misinformation During a Pandemic" (June 2020). NBER Working Paper No. w27417, Available at SSRN: <https://ssrn.com/abstract=3632644>

<sup>189</sup> Simonov, A. Sacher, S., Dube, J.P. and Biswas, S, "The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the COVID-19 Pandemic" (September 4, 2021). Columbia Business School Research Paper, Available at SSRN: <https://ssrn.com/abstract=3600088>.

<sup>190</sup> Ash, E., Galletta, S., Hangartner, D., Margalit, Y. and Pinna, M., "The Effect of Fox News on Health Behavior During COVID-19" (June 27, 2020). Available at SSRN: <https://ssrn.com/abstract=3636762>

<sup>191</sup> Pinna, Matteo and Picard, Léo and Goessmann, Christoph, Cable News and COVID-19 Vaccine Compliance (July 20, 2021). Available at SSRN: <https://ssrn.com/abstract=3890340>.

<sup>192</sup> *Ibid.*

<sup>193</sup> "Polio Spread Fueled by Vaccine Taboo" Fox News (Associated Press), May 4, 2005. <https://www.foxnews.com/story/polio-spread-fueled-by-vaccine-taboo>

nearly identical policy mandating vaccination or testing among its employees.<sup>194</sup> The owner of Fox News received the vaccine almost immediately after becoming eligible.<sup>195</sup>

Telling a blind person to step to the right so as to fall over a mountain cliff is unprotected speech.<sup>196</sup> Reflecting an externality, so too is an instruction to tell a sighted person to step into the path of another so as to trip that person over the cliff. The second order mortal externality renders the speech no less unprotected. Persuading an audience to forgo vaccination during a pandemic sets them up at the edge of a precipice without guardrails. That the unvaccinated then cross paths with others sending them over the precipice, without knowledge or consent, would not protect such speech even if it were policy relevant.

A key premise of government non-intervention is using the marketplace of ideas to resolve the merit of conflicting ideas.<sup>197</sup> But, in order for that market to function, it must have means of purging claims without merit. Like bad products artificially protected from competition, false claims have been artificially protected from consequence by the courts, crippling the functioning of and wasting literal lives and resources of free societies. Courts have failed to recognize that they have made the same mistake twice, once in antitrust by protecting inferior products rather than protecting markets for competition,<sup>198</sup> and again in free speech by protecting inferior speakers rather than the market for speech.<sup>199</sup> The solution, in both cases, is to allow the market (not government) to function properly and allow it to remove inferior products and false claims simply by allowing (not requiring) inferior sellers and deceitful speakers to suffer the consequences of their own advice and actions.

- 9) Free idea market solution: Stop protecting speakers at the expense of citizens and free markets for speech. Allow the market to clear false claims via standard contestation.

Deontologists have not held true to their position. If the reason for government non-intervention is to pass the test of an idea to the market, government should not intervene when the market decides an idea has failed the test. The paradox of antitrust jurisprudence

---

<sup>194</sup> Bauder, D. "Fox's Vaccine Criticism Focuses Attention on its Own Policy" AP News Sept 16, 2021. <https://apnews.com/article/joe-biden-business-health-arts-and-entertainment-fox-corp-26096a8781c7c7f1d6c0ddff98a5fe6d>

<sup>195</sup> Li, Kenneth "Murdoch receives COVID-19 vaccine as Fox News host casts suspicion on campaign," Reuters Dec 18, 2020. <https://www.reuters.com/article/us-health-coronavirus-murdoch/murdoch-receives-covid-19-vaccine-as-fox-news-host-casts-suspicion-on-campaign-idUSKBN28S2J7>

<sup>196</sup> On a list of twenty such examples, two more relevant cases include successfully encouraging a suicide and raising a false public alarm. Greenawalt, K. (1992). *Speech, crime, and the uses of language*. Oxford University Press.

<sup>197</sup> The "marketplace of ideas" has always described "ideas" not "facts," leading the law to treat facts and opinions differently. Facts are not copyrightable. Expressions with a modicum of originality are, implying an absence of a market in facts but the presence of a market for ideas. Blurring this distinction is ahistorical, incorrect, and "complicit in the erosion of the body politic," Waldman, A. (2017) "The Marketplace of Fake News," U. Pa. J. Const. L. pp 845-870.

<sup>198</sup> Bork (1978) **The Antitrust Paradox**. New York Free Press.

<sup>199</sup> Shiffrin calls recent jurisprudence "idolatry," overprotecting speakers at the expense of equality and of democracy. See Shiffrin, S. H. (2016). *What's Wrong with the First Amendment*. Cambridge University Press.



was that legal intervention intended to protect consumers and free trade markets artificially raised prices by protecting inefficient firms from consequences of competition.<sup>200</sup> The paradox of free speech jurisprudence is that intervention intended to protect citizens and free idea markets artificially raises harms by protecting those making false claims from the consequences of acting on those claims.

Where *Consumer Welfare* serves as a useful standard by which to judge competition in the market for products, *Universal Potentiation* might serve as such a standard to judge expression in the market for ideas. Importantly, it highlights what to protect, i.e. decision options and expression, and what not to protect, i.e. decision errors and harmful externalities.

In their concurrences to *Brandenburg v Ohio*, even the staunchest free speech absolutists, Justices Douglas and Black, took pains to defend and carve out the extreme speech exception of falsely shouting fire to cause a panic. The irony here is bitter. A putative news organization is falsely shouting “Not Fire!” when there is one, yet the courts are defending them after citizens get burned.

An obvious rejoinder is what then of personal agency? Would not the listener bear individual responsibility for his or her own choices apart from what the speaker recommends and are there not other sources of information? This has three answers.

First, it is the height of hypocrisy for a speaker to urge a course of action, here knowing the consequences, then assert they did not anticipate the result or that they bear no responsibility. In the present case, the defendant opposed a policy of intervention for others that it embraced for itself. The policy at issue, moreover, was not one of subjective fashion but one of objective science – was the pandemic a hoax or not, does social distancing work or not, are masks effective or not. Deontology in its pure form as expressed by Kant, not as practiced by the courts, not only requires telling the truth but reserves particular scorn for hypocrisy. No hypocrite can set their own actions contrary to those they apply to another and be practicing a universal rule. Hypocrisy lies at the heart of immorality under deontology. It is not defensible under the deontological principles espoused by the courts.

Second, in the context of multiple and conflicting data sources, holding the speaker accountable for population level decisions is challenging but possible. The hypothetical Alvarez case of providing false resume information to an employer is clear and simple. There is no competing source of misinformation and one party, the employer (or possibly a displaced worker), has a specific course of action based on the erroneous decision to hire. The case of Fox News misinformation is more challenging in that some other news organizations echoed Fox sentiments and that misinformation affects a population not just one entity. It is harder to assert Fox was decisive for a specific individual’s choices. Fortuitously, modern technology brings with it solutions to the problems that it enables.

---

<sup>200</sup> Ibid. Bork.

Data on audience choices show a preponderance of evidence, beyond a statistically valid reasonable doubt, that Fox News broadcasts were causal. Statisticians, for this specific case, have used Fox’s own audience as a control for competing news broadcasts, mobile phone data to track differences in lockdowns and social distancing among Fox News viewers, and consumer product data to track differences in use of hand sanitizers and masks. Assessment of population level responsibility is not only feasible but practical and more certain at precise levels of accuracy than most courts require. Indeed, such an assessment simply parallels “Market Share Liability” already applied in pharmaceutical cases where apportioning damages is decided by pharmaceutical market shares when who produced a specific generic drug is unknown.<sup>201</sup>

The third answer rejects the implied conclusion but embraces, in part, the stated assumption. The implied conclusion is to absolve the speaker of any responsibility whereas the assumption is that the listener bears all responsibility. One of Coase’s insights is that damage is often a function of choices made by *both* parties.<sup>202</sup> If this damage, a “negative good,” can be traded in the market, then assigning property rights can allow a market to trade its way back to efficiency. If, due to transaction costs, the negative good cannot be easily traded, then efficiency requires assigning liability to the lowest cost avoider. The reason is a simple problem of moral hazard that, freed of consequence, the lowest cost avoider has no incentive to reduce the damage. The alternative, in the Fox News case, is to assert that all audience members should first verify all broadcast claims – an absurdity of inefficiency. In this instance, as in the cases of air and water pollution, the lowest cost avoider is the party producing the misinformation and should therefore hold in principle and in principal the liability.

For purposes of epistemological clarity, I note that the *Paradox of Free Idea markets*, like Bork’s *Antitrust Paradox* is one of enforcement, not true logical contradiction. In both cases, courts have undermined their stated goals via protectionist intervention. Their enforcement actions have allowed bad products and bad ideas to circulate past expiration dates that competitive forces would have imposed, raising harms and reducing welfare for consumers and citizens alike. Deontological rules do, however, create true paradoxes of presence and of absence we must visit next.

## B. The Paradox of Individual Liberty

Standard models of decision theory overlook an important case where accurate policy information can itself induce decision errors. It is possible for one truth to cancel or obscure another truth. Signal jamming is a third form of dysfunctional information<sup>203</sup> causing harms

---

<sup>201</sup> See *Sindell v Abbott Laboratories*

<sup>202</sup> *Id.* Coase 1960.

<sup>203</sup> Another recognizable version is a “Denial of Service Attack” that floods a target with enough traffic to either bring it down or render it incapable of serving others. “Understanding Denial of Service Attacks” U.S. Cybersecurity & Infrastructure Agency. <https://www.cisa.gov/uscert/ncas/tips/ST04-015>. I thank Brendan Hemingway for this observation.

but it operates on a different principle. Rather than deceiving a decision maker with false news, or affecting third parties via externalities, it forces decision errors by denying access to the news one needs to know. It is the antithesis of idea diversity; it is noise; it is homogeneity. In terms of the Fundamental Welfare Theorem of economics, it implicates a third cause of failure in the market for ideas. After externality, and information asymmetry, it is monopoly.

---

*“What is the cost of lies? It’s not that we’ll mistake them for the truth. The real danger is that if we hear enough lies, then we no longer recognize the truth at all.” – Valery Legasov, 2019 (Chief of the Commission investigating Russia’s nuclear power disaster, from the movie Chernobyl)*

---

Information theory gives us the most rigorous proof of this phenomenon. Claude Shannon’s channel coding theorem formally shows that correcting errors is impossible once the flow of information exceeds a threshold determined by the noise level.<sup>204</sup> The theorem gives us the means of measuring the information volume present, defines the computer science “bit,” and forms

the basis of all telecommunications.

A human version of channel capacity forms the basis of Herbert Simon’s bounded rationality theory. “What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it.”<sup>205</sup> Explaining errors in human judgment that ensue from passing this boundary is one reason Simon received the Nobel Prize in economics. To more accurately explain behavior, he replaced a model of omniscient rational agents with one whose agents’ “capacities for rational action are limited.”<sup>206</sup> Other scholars have proven formally that once the volume of information hits the bounded rationality constraint, individuals’ information sets can become homogenized and associations of such individuals more polarized.<sup>207</sup> Thus a flood of Internet information can produce a form of cyberbalkanization. Algorithms that filter information to boost engagement and elicit strong emotional responses exacerbate this phenomenon.<sup>208</sup> Highlighting the inability to correct errors faster than they arrive, other scholars find “truth decay” occurs when people, no longer able to separate truth from noise, fall back on opinion and tribal beliefs to determine their choices.<sup>209</sup>

---

<sup>204</sup> Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423.

<sup>205</sup> Simon, Herbert A (1971). *Designing Organizations for an Information-rich World*. In M. Greenberger (Ed.), *Computers, Communications, and the Public Interest*. Baltimore, MD: Johns Hopkins University Press. pp. 37–52.

<sup>206</sup> <https://www.nobelprize.org/prizes/economic-sciences/1978/press-release/>

<sup>207</sup> Van Alstyne, M., & Brynjolfsson, E. (2005). Global village or cyber-balkans? Modeling and measuring the integration of electronic communities. *Management Science*, 51(6), 851-868.

<sup>208</sup> Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

<sup>209</sup> Kavanagh, J. & Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*. Rand Corporation.

Despite the theory and evidence from bounded rationality and from information theory, the Supreme Court has held steadfast to a policy that more speech is better speech. The landmark case *Buckley v. Valeo* struck down a “restriction on the amount of money a person or group can spend on political communication” because it “reduces the quantity of expression ... the number of issues discussed, the depth of their exploration, and the size of the audience reached.”<sup>210</sup> *Buckley v. Valeo* committed the Court to a position that more speech is necessarily better, deducing that government cannot therefore restrict volume. The stated, and reiterated, goal was “to secure, ‘the widest possible dissemination of information from diverse and antagonistic sources.’ ”<sup>211</sup> Decades later, the Supreme Court reaffirmed this position in *Citizens United*, writing “there is no such thing as too much speech”<sup>212</sup> and “it is our law and our tradition that more speech, not less, is the governing rule.”<sup>213</sup> As Kennedy observed for the majority, any “statute which chills speech can and must be invalidated.”<sup>214</sup>

These decisions are of a time and place that predate the manifest consequences of crowdsourced information. *Buckley v. Valeo* set policy in 1976 before the World Wide Web’s 1989 emergence, marked by the invention of hypertext.<sup>215</sup> The Supreme Court ruled on *Citizens United* in 2010, just four years after the 2006 founding of Twitter.<sup>216</sup> As of 2020, no free speech ruling by U.S. courts anticipated that a U.S. president might coax mob action to crowdsource the silencing of critics.<sup>217</sup> Courts have not recognized speech that cancels speech.

Yet, the Supreme Court has not always overlooked capacity constraints. When considering radio broadcast scarcity in 1969, it ruled in *Red Lion v. FCC* that: “No one has a First Amendment right to ... monopolize a radio frequency. ... It is the purpose of the First Amendment to preserve an uninhibited marketplace of ideas in which truth will ultimately prevail, rather than to countenance monopolization of that market, whether it be by the Government itself or a private licensee.”<sup>218</sup> Monopoly is a problem because the public has a right “to receive suitable access to social, political, esthetic, moral and other ideas.”<sup>219</sup> Interpreted through the lens of *Universal Potentiation*, an absence of idea diversity limits decision options due to incomplete information. *Red Lion* established the now disused

---

<sup>210</sup> *Buckley v. Valeo*, 424 U.S. 19 (1976).

<sup>211</sup> *Ibid.* Page 649, paragraph 60, citing *Op. Cit. NY Times v. Sullivan*, quoting *Associated Press v. United States*, 326 U. S. 1, 326 U. S. 20 (1945).

<sup>212</sup> Justice Stevens, dissenting. [Syllabus: \*Citizens United v. Federal Election Commission\*](#), Supreme Court of the United States. p. 83.

<sup>213</sup> *Op. cit. Citizens United v. FEC*, 558 U.S. 310, 361 (2010). P. 45.

<sup>214</sup> *Op. cit. Citizens United v. FEC*.

<sup>215</sup> “A Short History of the Web | CERN” <https://home.cern/science/computing/birth-web/short-history-web#:~:text=Tim%20Berners%2DLee%2C%20a%20British,and%20institutes%20around%20the%20world>.

<sup>216</sup> <https://en.wikipedia.org/wiki/Twitter>

<sup>217</sup> Wu, Tim (2018) “Is the First Amendment Obsolete?” *Michigan Law Review*, 117(3) p. 547-581.

<sup>218</sup> *Red Lion Broadcasting Co., Inc. v. FCC*, 395 U.S. 367 (1969), page 395.

<sup>219</sup> *Ibid.*

Fairness Doctrine, allowing political speakers broadcast access, on the basis of channel scarcity. The Fairness Doctrine still adheres to the more-is-better rule, creating more diversity by enabling more voices. Importantly, however, every channel has two capacity constraints, one at the transmission end and one at the receiving end. Courts have yet to recognize the latter.

By contrast, totalitarian regimes have recognized constraints at both ends and used both as weapons of censorship. Signal jamming has long been a common and deliberate form of censorship that countries use to prevent broadcast information reaching their citizens from adjacent countries.<sup>220</sup> During war, enemy combatants use signal jamming to prevent communication and coordination of troop divisions on the opposite side.<sup>221</sup> During peace, authoritarians now use signal jamming based on social media. “Rather than shutting down dissenting voices, these leaders have learned to harness the democratizing power of social media for their own purposes — jamming the signals, sowing confusion. They no longer need to silence the dissident shouting in the streets; they can use a megaphone to drown him out. Scholars have a name for this: censorship through noise.”<sup>222</sup> Rather than argue with dissent, authoritarians also use secretive operations to “distract the public and change the subject,” while using cheerleading to drown critics and solidify their hold on power.<sup>223</sup> U.S. politicians use social media to send censorial mobs after critics whose voices they wish to silence.<sup>224</sup> To implement “reverse censorship,” autocrats use “a sufficient volume of information to drown out disfavored speech.”<sup>225</sup>

Signal jamming and reverse censorship occur both as the monopoly speech of one voice and as the monopoly of many voices speaking as one. Public squares, print media, TV stations, and radio broadcasts represented the choke points of twentieth century market access. Digital public squares have removed choke points on transmission, only to replace them with choke points on reception.

Courts have not recognized this latter problem. Writing in defense of unlimited speech on the basis that more-is-better, the Supreme Court holds that: “Being free to engage in unlimited political expression subject to a ceiling on expenditures is like being free to drive an automobile as far and as often as one desires on a single tank of gasoline.”<sup>226</sup> No, this allegory is logically flawed, having missed or misunderstood the concept of congestion

---

<sup>220</sup> Jerome S. Berg (2008). *Broadcasting on the Short Waves, 1945 to Today*. McFarland

<sup>221</sup> [https://en.wikipedia.org/wiki/Radar\\_jamming\\_and\\_deception](https://en.wikipedia.org/wiki/Radar_jamming_and_deception)

<sup>222</sup> Coppins, McKay (2020) “The Billion-Dollar Disinformation Campaign to Reelect the President,” *The Atlantic*, March 2020. <https://www.theatlantic.com/magazine/archive/2020/03/the-2020-disinformation-war/605530/>

<sup>223</sup> King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American political science review*, 111(3), 484-501. P 484.

<sup>224</sup> Brooks, Rosa, “And Then the Breitbart Lynch Mob Came for Me,” *Foreign Policy*; Feb. 6, 2017. <https://foreignpolicy.com/2017/02/06/and-then-the-breitbart-lynch-mob-came-for-me-bannon-trolls-trump/>

<sup>225</sup> Ibid. Wu, p. 565.

<sup>226</sup> Ibid. Buckley, Footnote 18.

externalities. To use the judges' own metaphor, what these decisions have unleashed is the capacity to hire tens of thousands of drivers, jamming a city's streets, so that no other driver can deliver messages to their destinations. Congestion jamming *blocks* the widest dissemination of information from diverse sources. They have unleashed the social media equivalent of the denial-of-service attack.

The deontological paradox of presence is that unlimited amplification of one message means that no other message can be heard. It is a paradox of individual liberty taken to extreme. In order to protect liberty, the courts unwittingly undercut liberty. The deontological rule is that "protected" speech shall be protected absolutely. In the extreme, such a rule is self-canceling. To invalidate an absolute right, simply create a contradiction by pitting that right against itself in the manner that mathematicians proved there exist systems with no absolute truths. Consider the statement pair: {The next sentence is true. The previous sentence is false.} The inherent contradiction means the rule is void in at least one application. Not only must a given right balance *competing* rights – life, liberty, happiness – but that right must also balance *itself*. Illustrations prove instructive.

Consider an absolute right to life. Every society recognizes at least one limit on an absolute right to life: A sane person taking a life in the act of self-defense against attempted murder has an honorable expectation of absolution. Self-preservation is its own just end. The right to life is not so absolute that it cannot be taken in self-defense and the life that sought to extinguish another's right to life is the one more justly forfeit if only one survives.<sup>227</sup>

Consider an absolute right to anonymity. Suppose a criminal un.masks others' private identities yet seeks to avoid prosecution by using his right to anonymity as a means to hide from his crime. A reasonable conclusion is that the rights of the violator ought to be suspended precisely in order to protect the rights of everyone else. Failure to suspend the right to anonymity, in order to identify and prosecute the criminal, would be an illogical contradiction that would produce more frequent invasions of privacy. Enforcing the right would negate the reason for granting the right. Lower, not higher, levels of privacy would result.

A gun right that is used to take others guns, a right to life that is used to snuff others' lives, a speech right that is used to cancel others' speech – these cannot be justified. They are invalid. They are contradictions.

How then should resolution proceed? The nature of the contradiction yields the seeds of its own solution. The first seed grows from observing that signal jamming needs neither truth for fiction, only volume. Intervention can then proceed on the basis of reducing volume, independent of content. Like time, place, and manner restrictions, reducing speech volume, or adding friction to amplification, may proceed under intermediate rather than strict

---

<sup>227</sup> Van Alstyne, M. W. (2019). A response to fake news as a response to Citizens United. *Communications of the ACM*, 62(8), 26-29.

scrutiny.<sup>228</sup> Recognizing the nature of an externality problem, down-sampling a message already copiously repeated at the individual level or adding a progressive ad tax at the expenditure level, might serve to reduce congestion in communication channels. Friction has almost no effect on awareness of a message already monopolizing a market.

The second seed identifies the party that first abuses their right by pitting it against the rights of others as the party whose right merits restriction. Other parties affected by abuse but not instigating conflict should have their rights protected if for no other reason than that the alternative invites abuse. No act is dissuaded if a right cannot be forfeit purely for exercising that right in the extreme. If the stated goal of the deontological view is the greatest preservation of rights, then reason requires that right be restricted, when pitted against itself, precisely for that party who seeks to revoke the rights of others.

The principle of *Universal Potentiation* is to expand an expressive right of reach until it can go no further. What then is the boundary on reach? It is the point where one person's expression prevents that of another. Thus, we have:

- 10) Individual liberty solution: Ex ante, the only justifiable restrictions on expressive reach are those to safeguard the reach rights of others. Protect individual liberty by checking only monopoly. Add friction to amplification that prevents others from being heard.

### C. The Paradox of Preventing Government Tyranny

The forgoing analysis proceeds from an excess in exercise of free expression rather than from rights suppression. Enriching individual liberty, however, is not how the First Amendment works. Instead, it acts as a negative right, not a positive one. It operates not by raising the hands of its citizens but by tying the hands of its government, an absence of power not a presence. Does the forgoing analysis apply when the mechanism operates in reverse? As free expression advocates proclaim, the imbalance in power between state and individual is so great that only by drawing a bright line circumscribing state action can laws balance freedom and authority.<sup>229</sup> Limiting power limits arbitrary use of power. Thus, to prevent state tyranny, with only a handful of exceptions, the deontological rule limits state intervention regardless of consequence. Yet, as before, the paradox of the deontological approach is that taken to extremes it is self-cancelling. In the name of preventing state tyranny, it results in the onset of that very tyranny.

How might this be so? How might a right denied be pitted against itself so as to render a contradiction? As a conclusion of formal logic, if we have  $P \rightarrow Q$ , then we must also have  $\sim Q \rightarrow \sim P$ . Thus, if we can construct a contradiction based on a positive assertion, we ought to be able to construct a contradiction based on a negative assertion. Proceeding as follows, note that population of  $N$  people can form at most  $2^N$  subgroups and therefore  $2^N$  possible

---

<sup>228</sup> Miller, Erin. (2021) "Amplified Speech," *Cardozo Law Review* 43(1), P 1-69.

<sup>229</sup> *Ibid*, Emerson 1962.

governance mechanisms.<sup>230</sup> Call each of these modes of governance a religion or ideology. Let one of these represent a cancerous ideology that seeks to suppress all other  $2^N-1$  ideologies. Assuming all  $2^N-1$  others adhere to the absolute non-intervention rule, but the cancerous ideology intervenes, the end result is self-cancellation of the original non-intervention rule for all groups that adhere to it. Enforcing it leads to its complete revocation and non-enforcement. A deontological approach that holds the rule as absolute over a population is per se an absolute failure. Taken to extreme, it cannot be consistent with itself.

This self-contradiction, based on a right denied, illustrates the irony that tolerance is a virtue up to but not past the point of tolerance of intolerance, where it becomes a vice. To tolerate the spread of intolerance produces the abnegation of tolerance. Derivation of the *Paradox of Preventing Government Tyranny* restates, more formally, Popper's *Paradox of Tolerance*, which he shows is equally illogical in the extreme:<sup>231</sup>

*Unlimited tolerance must lead to the disappearance of tolerance. If we extend unlimited tolerance even to those who are intolerant, if we are not prepared to defend a tolerant society against the onslaught of the intolerant, then the tolerant will be destroyed, and tolerance with them. — In this formulation, I do not imply ... that we should always suppress the utterance of intolerant philosophies; as long as we can counter them by rational argument and keep them in check by public opinion, suppression would certainly be most unwise. But we should claim the **right** to suppress them if necessary even by force; for it may easily turn out that they are not prepared to meet us on the level of rational argument, but begin by denouncing all argument; they may forbid their followers to listen to rational argument, because it is deceptive, and teach them to answer arguments by the use of their fists or pistols. We should therefore claim, in the name of tolerance, the right not to tolerate the intolerant. We should claim that any movement preaching intolerance places itself outside the law and we should consider incitement to intolerance and persecution as criminal, in the same way as we should consider incitement to murder, or to kidnapping, or to the revival of the slave trade, as criminal. — Karl Popper 1945<sup>232</sup>*

---

<sup>230</sup> Mathematicians might assert that one of these subgroups is simply the empty set. Political scientists might then observe that the empty set simply implies anarchy.

<sup>231</sup> For nomenclature, I have chosen to retain the *Paradox of Preventing Government Tyranny* in preference to the terminology *Paradox of Tolerance* in order to focus on the mechanism, governance, that represents the point of intervention. Free speech writing, since at least the time of Mill (1859), has focused on tying the hands of the state in order to prevent tyranny. The question then is specifically when to untie state hands in order that a worse state might be avoided as distinct from when to revoke tolerance in order that intolerance might be avoided.

<sup>232</sup> Popper, Karl (1994) "The Open Society and Its Enemies" Princeton University Press. Fn 4. P 581.



This contradiction brings with it a second false negative in the Brandenburg test, hinging on incitement to imminent lawless action and propensity to do so. The usurpation of legitimate power by tyrannical power fits no previously enumerated category and has potential to proceed via legal means. The challenge of lawless tyrannical governance is obvious when posed by an enemy state. Although the Supreme Court rendered a different logic, the principle of preserving legitimate governance can fully justify suppressing “publication of the sailing dates of transports or the number and location of troops” during wartime.<sup>233</sup> The challenge of lawful tyrannical governance is less obvious, however, when posed by citizens from within. Elaboration of the potential to proceed to tyranny via legal means is documented by those who have done so:

*“We enter parliament in order to supply ourselves, in the arsenal of democracy, with its own weapons... If democracy is so stupid as to give us free tickets and salaries for this disservice, that is its affair. ..We do not come as friends, nor even as neutrals. We come as enemies: As the wolf bursts into the flock, so we come.” – Joseph Goebbels 1928 in Der Angriff*

[xxx Trump’s Big Lie]

What happens when the lie defines what is legal? What happens when those who believe the lie use legal means to appoint those who decide the lie? At which point in time was illegality imminent? At which step in the chain were legal actions illegal?

Neither the action nor the outcome is illegal to one who believes the lie. Its effects are both legal and moral in the eyes of the credulous and if that lie metastasizes, becoming nourished and entrenched by repetition, it infects others still. Their rights being equal to those of others, their beliefs of no lesser weight, the injustice of usurping sovereign authority becomes complete once their decisions bind. Critically, it is not necessary for the credulous to be a majority, only that they hold positions of lawful authority.

How might we know that a lie rises to such a dangerous level? When is there reason to act? One test is whether failure to act invites the very injustice that inaction sought to avoid. As with the *Paradox of Individual Liberty*, the nature of the contradiction is cue to its resolution. The test then is when to untie state hands in order that a worse state might be avoided. This condition establishes necessity to act in order to escape the paradox.

A second test, respecting the right to exercise authority, is that the liar rejects any outcome limiting that authority arrived at via neutral decision processes, accepting only those outcomes it alone has power to certify. It proves its intolerance, rejecting impartial juries and rejecting due process. It seeks to absorb that state function designed for adjudicating fair outcomes behind a veil of ignorance. With this power, the liar may climb from the minor rungs on the ladder of state to the summit of its operation. This mechanism permits

---

<sup>233</sup> *Near v. Minnesota* 283 U.S. 697 (1931).

the usurpation of authority and with it the usurpation of others' sovereignty. This test seeks to identify partisan certification and rectify it.

A third test invokes *Universal Potentiation*. Do the teleological implications of suppressing the claim bind a *successor* government, not those in power, to a condition of at least as much justice and as many decision options for citizens as bind the current institution? Has the expressive power of citizens and their ability to make unencumbered decisions endured or, better, risen? The third condition assures forward progress in implementing just governance. It allows for stable social change, passing from one government to another, adhering to and improving upon a principle of just governance.

What then are the choices, the avenues out of this injustice? The first, and usual default, is inaction. State non-interference is what courts have assumed, allowing events to unfold on the premise that truth has the power "to get itself accepted in the competition of the market".<sup>234</sup> It is paradoxically the tolerance that replaces itself with intolerance when a market has seriously corrupted sellers. A second intervention might seek to change the beliefs of either side. The power to change beliefs, except by reasoned persuasion, would be a dangerous tool in any hands. Any state granted such power could create its own Orwellian reality. A third intervention might be to accept the claim but adjudicate the injustice *ex post*. The conundrum is that, having usurped power, those who would have power have every reason to maintain it and no reason to admit obtaining it based on a lie. A fourth intervention is violence. Governance mechanisms having failed, extra legal means become the instruments of social change. Despite advocating openness and tolerance, even Popper allows force as a last resort. Yet, a fifth alternative is to limit the spread of the lie *ex ante* before it metastasizes.

*None* of these options are attractive. The point in their enumeration is that the first four all lead to tyranny; only the fifth avoids it. Other options might exist but, no matter their number, the conclusion remains intact so long as alternatives maintain, enable, or simply fail to overturn unjust governance once established.

The implication is that government intervention to chill lies is not only warranted but mandated when the effect of the lie's action is to produce unjust governance. Failure to act reifies the very injustice that inaction sought to avoid.

The principle of *Universal Potentiation* is to expand a right of free expressive until it can go no further. What then is the boundary on expression? It is the point where one person's expression violates the rights of another or violates the governance that protects those rights. Thus, we have:

- 11) Preventing tyranny solution: *Ex ante*, the only justifiable restrictions on expression are those to prevent violence to the rights of others or to prevent violence to the system of

---

<sup>234</sup> *Abrams v. United States* (1919), Holmes dissenting.

governance that safeguards those rights. Reject such tyrannical speech as would itself lead to tyranny. For the prevailing government, behind a veil of ignorance, apply the criteria of a non-partisan *successor* government that would grant no lesser liberties than those present.

Across the set of utilitarian responses to deontological problems there have only been three cases and each offers a different prescription for the state's role in free speech. The first, the *Paradox of Free Idea Markets* exhibits no true paradox, only one of enforcement that undermines its own goals. From a government perspective, no *ex ante* intervention is warranted. Indeed, the solution is to step aside more often in order that the market might clear itself of false claims by internalizing harms. The second, the *Paradox of Individual Liberty*, is a true deontological paradox of presence where the expansion of one party's right eclipses that same right in another. From a government perspective, *ex ante* intervention is warranted but not content based. Action serves only to turn down the volume on one deafening voice in order to bridle suppression of others' voices. The third, the *Paradox of Preventing Government Tyranny*, is a true deontological paradox of absence. Compelling inaction to forestall tyranny enables the actions of those who would install tyranny. From a government perspective, *ex ante* intervention is warranted yet a principle holds that the content restrictions are so narrow, they serve only to ensure the legitimacy of a successor state.

## X. Tests & Objections

The complexity of intervening in news streams means that balancing different objectives inevitably breaks some constraint. This section seeks to address the most common objections to interventions in any market for free speech. These four are among the most common.

### 1) Platforms do not produce the content they propagate.

If social media platforms do not author the content that cause harm, why should having them internalize these costs be more efficient than having content creators bear these costs? Section 230 grants them immunity on this basis. A more robust analysis provides two answers based on governance and transaction costs. First, social media platforms already internalize the positive externalities of social networks; they need only internalize the negative externalities as well. Social media platforms are built on and derive their power from network effects. Their purpose is to foster connections.<sup>235</sup> The contacts and activity of one user benefit other users. These are externalities and when they are positive the platform already encourages them and profits from them by interposing itself and monetizing ads

---

<sup>235</sup> Facebook's original mission statement was "Making the world more open and connected." Constine, J. "Facebook changes mission statement to 'Bring the world closer together'" TechCrunch <https://techcrunch.com/2017/06/22/bring-the-world-closer-together/>

between connections. When harms occur on-platform, as in the case of harassment or fraudulent products, the platform already addresses them. Platforms only need motivation to take actions they already take yet must do so for harms that occur off-platform in addition to those that occur on-platform.

Although social media platforms do not author the content they dispense, any claim they exert no influence over members is disingenuous. They actively engage in orchestration. They are the governments of their ecosystems with authority to regulate participation, prices, competition, and intellectual property within their regimes.<sup>236</sup> The venture capitalists who invest in platforms, not merely the economists, have stated as much.<sup>237</sup> When the citizens of one country suffer the pollution of another, the government of the former might reasonably negotiate with the government of the latter, especially when polluters in the latter are invisible to citizens of the former. Although neither government itself produced the pollution, lax rules in the source country are at least partially responsible for pollution in the harmed country. In this case, the costs of harmed individuals bargaining with each polluting firm greatly exceed those of collective bargaining. This argues for negotiating with the government of the polluting country as the highest leverage point of intervention.

Second, platforms do, in fact, represent the nexus of lowest transaction costs.<sup>238</sup> Social media platforms, unlike one-way broadcast media, orchestrate the activities of their users. In order to facilitate membership and engagement, they reduce friction on participation and production. They provide tools for creation, tools for consumption, and feedback on impact. In fact, no party has greater visibility than social media platforms into the nature of misinformation transactions. Without information supplied by the platforms, not even the authors themselves know who has shared or who has read their campaigns. Transaction cost economics weigh in favor of intervening at the point of lowest cost and greatest transparency, in this case, the point of the platform. Social media platforms represent the point of greatest leverage.

## 2) How can this or any mechanism decide what is really true?

The effort to ascertain truth has two approaches, one practical and one theoretical. In practice, courts routinely grapple with the question of whether a claim is true "beyond a reasonable doubt." They address this challenge by raising the level of threshold to the import of the problem. For example, the legal tests that we apply to free speech laws and their breach must pass three different thresholds. Roughly stated, these are *rational basis review*: does the law relate to a legitimate end and has it in truth been violated? This dispenses with specious cases. A higher level is *intermediate scrutiny*: does violating the law affect a basic

---

<sup>236</sup> Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform revolution*. Norton. Chapter 8 - Governance. Tirole, J. (2017). *Economics for the common good*. Princeton University Press.

<sup>237</sup> Brad Burnham, "Web Services as Governments," Union Square Ventures, June 10, 2010, <https://www.usv.com/blog/web-services-as-governments>.

<sup>238</sup> Juliet Yu, Alibaba. Munger, M. (2015). "Coase and the Sharing Economy," in *Forever contemporary: the economics of Ronald Coase*. Veljanovski, Cento (ed). Institute of Economic Affairs, pp 187-208.

right and has it in truth been violated? This protects individual rights. The highest level is *strict scrutiny*: intervention must protect a compelling government interest and be narrowly tailored. This protects group rights. Similarly, the laws around "duty of care" for product liability are particularly vague and differ by state in the same manner that free speech laws differ by country. And yet, as a practical matter, we deal with them. We apply local context and change the threshold for certainty according to the severity of the decision.

As a theoretical matter, we cannot know absolute truth. This objection simply re-asks Hilbert's "*Entscheidungsproblem*" in a new context. The Church-Turing thesis tells us that certain statements cannot be proven true or false. A more precise statement of Hilbert's decision problem, grounded in logic and philosophy, is as follows. Given a system of claims, is it possible to definitively prove the collected assertions are true? The answer, in general, is no. Posed in 1928, this hard question was not answered until 1936, when Alonzo Church and Alan Turing independently developed methods to prove that an infinity of claims are undecidable. Modern computer theory provides an interesting clarification. If statements are made at one end of a communication channel, can one be certain that identical statements are received at the other end of that channel? In effect, data corruption – literally false news introduced into the channel – can be repaired using error correcting methods but only up to a point. Shannon's Channel Coding Theorem, which forms the basis of all modern communications, proves that arbitrarily small error in communicating a fact is not achievable. Error correction is only possible up to a fixed and finite boundary. It is impossible past that boundary.

Thus, the question presupposes an answer that cannot be given. This objection is used to dispatch any approach that cannot solve the problem, which is an unfair critique because *no* approach can solve the problem. We can only know truth to a given number of bits. Interestingly, this comports well with the practical solution of accepting a claim as true "beyond a reasonable doubt."

- 3) Edge cases between true and false invalidate the mechanism. Whatever the boundary conditions, it is always possible to split a boundary with careful wording.

The existence of an edge case is not a legitimate challenge to any governance mechanism generally, let alone a fair news mechanism specifically. There does not exist any useful mechanism for which there do not exist edge cases between true and false. The decision criterion "Always Guilty" has no edge cases but also no practical application, as does its opposite "Always Innocent". Even our most cherished and most absolute rights have edge cases. Is the right of free speech absolute? We admit slander, libel, and incitement to violence as exceptions. Is the right to life absolute? We admit self-defense as an exception. The existence of edge cases can be used to exclude every mechanism, which leaves only the null mechanism, thus it fails as a legitimate test. Rather, the test should be whether one mechanism adjudicates edge cases better than the alternative. Importantly, the alternative is not the null set of no mechanism at all. In our case, the alternative is the present mechanism being used by social media platforms and, judged in terms of efficacy

and absence of bias, those leave much to be desired. The proper challenge is therefore to articulate the alternative mechanism and show why it does better than the proposal under consideration. Across a weighted sum of false positives and false negatives, which rule achieves more social value? Admittedly, this is a high bar. The best challenge is a superior mechanism design. The best and most challenging objection is thus an act of creation and not merely an act of rejection.

- 4) Reasonable people will disagree and those who dislike a decision simply seek to discredit the decision maker. Unless adjudication is indisputably impartial, partisans will not accept results.<sup>239</sup> We cannot avoid the problem of who gets to decide.

Conservatives may reject a decision whose outcome favors a liberal view. Liberals may reject a verdict whose outcome favors a conservative view. This objection raises separate issues of reconciling opposing views and of decision legitimacy. On the issue of reconciling conflicting views, there are three reasons why *requiring* agreement is ill advised.

- i. People do not universally wish to be convinced of a position nor do they grant third parties the moral authority to convince them. They often reject data that disagrees with their identity or world view or position in life. "It is difficult to get a man to understand something when his salary depends on him not understanding it."<sup>240</sup> More carefully, one may reasonably ask, what gives the mechanism designer the moral authority to assert the righteousness or truthfulness of the mechanism's verdict?<sup>241</sup> Absent such authority, perhaps the empowered view should shift its position to the disempowered view.
- ii. Mechanisms that *require* agreement cause moral hazard. If partisans know they will be bought out, with resources needed to convince them, they can exaggerate their claimed protests and supposed beliefs. These beliefs are not themselves verifiable. The social cost to overcome this mechanism-induced moral hazard could be in excess of the value of the verdict, producing social waste. The alternative, coercion, risks reaching the wrong conclusion simply by placing the power of coercion with one or another party.
- iii. The most compelling reason, however, that requiring agreement is not a valid test is an artifact, again, of the *Entscheidungsproblem*: knowing absolute truth is absolutely impossible. If one *unbiased* party cannot know or even communicate certain truth, it is pointless to require multiple *biased* parties to agree on that truth. Universal agreement is an impossible standard.

If reconciliation is infeasible, then in what sense might a verdict be legitimate? The solution is one we recognize in other contexts as procedural fairness.<sup>242</sup> Partisans must agree *ex ante*

---

<sup>239</sup> This objection arose in a conversation with the misinformation team at Facebook.

<sup>240</sup> Sinclair, Upton. (1994). *I, Candidate for Governor*. University of California Press.

<sup>241</sup> Mill, J. S. (1966). On liberty. In *A selection of his works* (pp. 1-147). Palgrave, London.

<sup>242</sup> Van den Bos, K., Wilke, H. A., & Lind, E. A. (1998). When do we need procedural fairness? The role of trust in authority. *Journal of Personality and Social Psychology*, 75(6), 1449.

to the method for deciding what's true, then commit to abide by the impartially administered verdict.

Thus, to operate a market for truth, we can rely on established administrative practices that we already use for trust and legitimacy. Taking our own government as precedent, consider a design where we split fake news oversight into legislative, judicial, and executive offices. A legislative body gets to define "fake ads." Despite their differences, even Fox News, CNN, and the New York Times might be able to agree on a working definition of fake news independent of specific cases. A judicial body gets to decide whether a specific case represents an instance of fake news according to this definition. Fact-checking organizations or juries of peers might play this role only now they must judge according to the definition provided by the legislative body. Jurors do not get to use their own individual definitions. Finally, the executive branch enforces these definitions and decisions. Social media platforms like Facebook and Twitter can play this role but they decide neither the definitions nor the outcomes of challenges. By dividing the branches of fake news governance, we recreate an institution where no branch judges truth as applied to itself and no branch has an economic incentive to bias its behavior to get rich. The divided process should therefore be free of conflict of interest, less biased due to random sampling, and by design more legitimate.

## XI. Conclusions

Societies cannot function that cannot agree on facts. Is the planet warming or not? Do pandemic vaccines work or not? Is the president in office legitimate or not? Twenty-first century fake news has pushed the boundaries of our institutions past the boundaries of our twentieth-century frameworks.

The goal of this article has been to reshape our understanding of the problem in order that we might use different tools to carve better answers. One step in this process is to shift the framing from a deontological perspective to a utilitarian perspective. Our courts' current use of deontological categories lead, in the extreme, to both paradoxes of implementation and of logic whereby outcomes undermine goals. Three paradoxes arise from absolutist free speech jurisprudence: in protecting markets of ideas, it undermines them, in protecting individual liberties, it defeats them, and in protecting against government tyrannies, it embraces them.

By contrast, use of a utilitarian perspective allows us to not only escape these paradoxes but also deploy well-sharpened tools from decision theory, information asymmetry, externality economics, and mechanism design. Pragmatism also admits tools from computer science with its more precise models of information, errors in transmission, and limits to comprehension.

---

<sup>243</sup> Nat'l Federation of Indep. Businesses et. al. v Sebelius, Secretary of Health and Human Services et. al.

Within this set, the first design options draw from choice architecture. Better designs reverse amplification of dysfunctional information and seek to place the burden for disseminating falsehoods on the author rather than the medium or the reader. They also go beyond traditional forms of transparency in order that parties harmed by misinformation can do more than learn of the harm; they can act to reverse it. Information together with access provides the means for injured parties to act so as to fulfill their own potential.

Recognizing dysfunctional information as a pollution problem summons a second set of tools – those of Pigou and of Coase – to address externalities. These grant the power to force those who disseminate damage to internalize that damage, overturning the incentive to propagate falsehoods for private gain at great social cost. These also give us means to assign liability to platforms for their share of harms. We can thus revise Section 230 in ways that scale with the volume of content to any level of certainty we desire. These tools can reform the business models of the twenty-first century information giants in the manner they reformed the pollution habits of the twentieth century industrial giants.

Adding the tools of information economics makes possible signaling and screening in a Market for Truth. A primary advantage is that the state need not regulate content at all. Informed authors, who know they are honest, can credibly signal integrity of their claims. Uninformed readers can then distinguish dishonest authors who are unwilling to warrant their dishonest claims. A secondary advantage is a change in cost structure. To date, no other mechanism has addressed the integrity handicap that producing fake news costs less than producing true news. These tools can also reconcile the inconsistent laws in paid advertising that have afflicted print versus broadcast versus internet media. Such laws arose from times and places past and have long needed reform.

A final tool, a principle of *Universal Potentiation* seeks to provide a standard by which to measure free speech protection. The principle supports individual liberty, offers properties parallel to those of the First Fundamental Welfare Theorem of Economics, and is at least consistent with the free speech goals of seeking truth, providing autonomy, participating in self-governance, and enabling stable social change. Grounded in decision theory on the premise of creating decision options that favor speakers, it shifts the focus of performance from protecting categories to protecting value as determined by decisions implicated and decisions changed. On this basis, value created in wise choices and value destroyed through errors and externalities can both be assessed from speaker and listener vantage points.

This utilitarian response to deontological problems has three implications. First, much false information passing through idea markets can be swept away if courts would but step aside. Repeating their mistake in antitrust in the context of free speech, they have protected speakers at the expense of speech markets as they formerly protected competitors at the expense of competitive markets. This *Free Speech Market Paradox* closely resembles Bork's *Antitrust Paradox*. With the means to measure value comes the means to measure harms and thus the means to hold those who cause harms accountable. In this context, government regulation of content is unwarranted.



A second implication is to introduce a different form of decision error based on dysfunctional information that has little recognition in current law. Signal jamming arises when so much information floods an idea market that faults necessarily occur. Either requisite information is missing, having been crowded out, or volume exceeds channel capacity, so error correction is impossible. Decisions suffer due to incomplete information or error introduction. The *Paradox of Individual Liberty* is that one person's volume cancels another person's voice. A problem of idea market monopoly then arises in two forms. At one extreme, monopoly is one voice amplified a million times as law has lifted speech volume limits (*Buckley v Valeo, Citizens United v FCC*). At the opposite extreme, it is a million voices amplifying but one idea seeking to drown differences in opinion. The latter case was not possible pre-internet platform.<sup>244</sup> In this context, government regulation of content is also unnecessary. Instead, intervention need only reduce volume of the monopolizing voice sufficient to allow diversity of ideas or rather to allow error correction rate to match volume flow rate.

The third implication is that government intervention to curb dysfunctional lies *is* warranted when the rule of law is itself at stake. Current legal practice gives rise to a *Paradox of Preventing Government Tyranny*. Tying the hands of government to prevent abuse of power unties the hands of immoral agents who lie to the credulous, granting them means to seize authority and thus abuse their power. A regime of absolute tolerance of intolerance replaces itself with a regime of intolerance.<sup>245</sup> Then, government regulation of content is warranted but is so narrowly prescribed as to ensure legitimacy of a successor government.

This is a work-in-progress distributed to solicit feedback. Pls send comments, critiques, objections, or encouragement to [mva@bu.edu](mailto:mva@bu.edu). Thank you.

---

<sup>244</sup> Ibid Wu. Ibid Miller.

<sup>245</sup> Ibid Popper.