

# University of Virginia School of Law

Public Law and Legal Theory Research Paper Series 2022-61

Law and Economics Research Paper Series 2022-19



SCHOOL of LAW

September 2022

## Truth Bounties: A Market Solution to Fake News

By

Yonathan A. Arbel  
University of Alabama School of Law

Michael D. Gilbert  
University of Virginia School of Law

## University of Alabama School of Law



Working Paper No. 4204862

Abstract 4204862

A complete index of University of Virginia School of Law research papers is available at: Law and Economics:

<http://www.ssrn.com/link/U-Virginia-LEC.html>

Public Law and Legal Theory: <http://www.ssrn.com/link/U-Virginia-PUB.html>

# TRUTH BOUNTIES: A MARKET SOLUTION TO FAKE NEWS

*Yonathan A. Arbel\** & *Michael D. Gilbert\*\**

*102 North Carolina Law Review* (Forthcoming)

## ABSTRACT

False information poses a threat to individuals, groups, and society. Many people struggle to judge the veracity of the information around them, whether that information travels through newspapers, talk radio, TV, or Twitter. Concerned with the spread of misinformation and harmful falsehoods, much of the policy, popular, and scholarly conversation today revolves around proposals to expand the regulation of individuals, platforms, and the media. While more regulation may seem inevitable, it faces constitutional and political hurdles. Furthermore, regulation can have undesirable side effects and be ripe for abuse by powerful actors, public and private.

This Article presents an alternative for fighting misinformation that avoids many pitfalls of regulation: truth bounties. We develop a contractual mechanism that would enable individuals, media, and others to pledge money to support the credibility of their communications. Any person could claim the bounty by presenting evidence of the falsity of the communication before a dedicated body of private arbitrators. Under the system we envision, anyone consuming information on the internet would know immediately if a given communication had a bounty attached, whether the communication had been challenged, and whether the challenge succeeded or failed. As John Stuart Mill recognized, we can trust our grasp of the truth only by putting it to the fire of challenge. Truth bounties open the challenge to all.

---

\* Associate Professor of Law, University of Alabama.

\*\* Vice Dean and Professor of Law, University of Virginia. For helpful comments, we thank Liam Bourque, Danielle Citron, Shahar Dillbary, Benjamin McMichael, and participants in the Harvard Law School Law & Economics Seminar, the American Law & Economics Conference at Columbia University, the Midwest Law and Economics Conference at the University of Chicago, the Maryland Carey Law Virtual Constitutional Law and Economics Workshop, and a seminar at Université Paris Panthéon-Assas. Alex Wilson, Gilberto Gomez, and Boston Topping provided diligent research assistance and offered many thoughtful suggestions.

<i>Introduction</i> .....	3
<i>I. Misinformation and Proposals for Reform</i> .....	9
A. Increasing the Supply of True Information .....	10
B. Decreasing the Supply of False Information .....	12
C. Distinguishing True and False Information .....	16
<i>II. Fake News in the Marketplace of Ideas</i> .....	17
A. The Spillover Problem .....	18
B. Information Asymmetry .....	22
<i>III. Truth Bounties</i> .....	30
A. Skin in the Game .....	30
B. The Voluntary Pledge .....	32
C. Challenges and Fees .....	34
D. Arbitration .....	35
E. Rewards and Signals.....	38
F. Equilibrium and Use Cases .....	40
G. Sustaining Truth Bounties .....	46
<i>IV. Bounties and the Freedom of Speech</i> .....	47
A. Contracts vs. Torts.....	47
B. Equity and Access.....	49
C. Hands-On, Hands-Off, and the Invisible Hand .....	51
<i>Conclusion</i> .....	53

## INTRODUCTION

False information threatens society.<sup>1</sup> Many people struggle to judge the veracity of the information around them, whether that information travels through newspapers, talk radio, or Twitter. Recent allegations about bad vaccines, stolen elections, and sex crimes by politicians demonstrate the problem.<sup>2</sup> With social media's speed and the amplification of content optimized for likes, clicks, and shares rather than value, and sometimes at the behest of foreign powers, the mixture of truth and lies churns.<sup>3</sup> Some fake stories take hold, driving opinions, trends, and possibly elections. Alarmed, leading scholars have turned to this issue with a sense of urgency, offering a menu of regulatory reforms in books,<sup>4</sup> law reviews,<sup>5</sup> conferences,<sup>6</sup> and the popular press.<sup>7</sup>

The sense of urgency notwithstanding, and although “fake news” consumes popular discussion, the basic issue is not new.<sup>8</sup> Beyond mass media and the press, in fields as varied as business, law, medicine, and politics, lying can be beneficial, so people have an incentive to lie. When lying can be legally or socially punished—fines, imprisonment, a loss of office or reputation—it can be deterred, at least in part.<sup>9</sup> This explains why products like toasters and providers like doctors mostly perform as advertised. If they did otherwise, people could

<sup>1</sup> See, e.g., Robert Chesney & Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CAL. L. REV. 1753 (2019).

<sup>2</sup> See, e.g., Anthony J. Gaughan, *Illiberal Democracy: The Toxic Mix of Fake News, Hyperpolarization, and Partisan Election Administration*, 12 DUKE J. CONST. L. & PUB. POL'Y 57, 74 (2017) (“The ease with which fake news, misinformation, and false allegations spread like wildfire is now a disturbing hallmark of modern politics.”).

<sup>3</sup> Brendan Nyhan argues that concerns with fake news reflect a “moral panic,” with little systematic evidence “to demonstrate that the prevalence of misperceptions today (while worrisome) is worse than in the past.” Brendan Nyhan, *Facts and Myths about Misperceptions*, 34 J. ECON. PERSP. 232-33 (2020). Similarly, Yochai Benkler and others contend that misinformation is a real concern but is elite-driven, rather than grassroots social media phenomenon. Yochai Benkler et al., *Mail-In Voter Fraud: Anatomy of a Disinformation Campaign*, 6 BERKMAN CTR. RSCH. PUB. 1 (2020).

<sup>4</sup> MARTHA MINOW, *SAVING THE NEWS* (2021); RICHARD L. HASEN, *CHEAP SPEECH: HOW DISINFORMATION POISONS OUR POLITICS—AND HOW TO CURE IT* (2022).

<sup>5</sup> Cass R. Sunstein, *Falsehoods and The First Amendment*, 33 HARV. J. L. & TECH. 388 (2020); Allison Orr Larsen, *Constitutional Law in an Age of Alternative Facts*, 93 N.Y.U. L. REV. 175 (2018); Abby K. Wood & Ann M. Ravel, *Fool Me Once: Regulating Fake News and Other Online Advertising*, 91 S. CAL. L. REV. 1223 (2018); Ari Ezra Waldman, *The Marketplace of Fake News*, 20 U. PA. J. CONST. L. 845 (2018); Nabihah Syed, *Real Talk about Fake News: Towards a Better Theory for Platform Governance*, 127 YALE L.J. F. 337 (2017-2018); Madeline Lamo & Ryan Calo, *Regulating Bot Speech*, 66 UCLA L. REV. 988 (2019); Erwin Chemerinsky, *False Speech and the First Amendment*, 71 OKLA. L. REV. 1 (2018).

<sup>6</sup> See, e.g., *International Conference on Media Manipulation, Fake News and Disinformation*, World Academy of Science, Engineering, and Technology, <https://waset.org/media-manipulation-fake-news-and-disinformation-conference> (last visited June 19, 2022); *Fighting Fake News Workshop*, Yale Law School Information Society Project, <https://law.yale.edu/isp/initiatives/floyd-abrams-institute-freedom-expression/practitioner-scholar-conferences-first-amendment-topics/fighting-fake-news-workshop> (last visited June 19, 2022).

<sup>7</sup> Ellen Maloney, *Professor Examines How Social Media Incites Spread of Fake News*, THE DAILY FREE PRESS (Oct. 5, 2020); Kevin Aslett, *It's Not Easy for Ordinary Citizens to Identify Fake News*, THE WASHINGTON POST, (Apr. 7, 2020).

<sup>8</sup> See, e.g., Marshall W. Van Alstyne, *Free Speech, Platforms, and the Fake News Problem*, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3997980](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3997980). (recounting how frescos from the 13<sup>th</sup> century BC record falsely the victory of Ramses over the Hittites).

<sup>9</sup> Part of the problem is that not all lies are detectable, a systemic issue in the market for “credence goods”—goods whose utility is not apparent even after consumption, such as vitamins, prayer, legal advice, etc. See J. Shaha Dillibary, *Trademarks as a Media for False Advertising*, 31 CARDOZO L. REV. 327, 341 (2009).

sue for the harms they suffer, the state could regulate, reputations would suffer, and profits would shrink. In many domains, economic, legal, and social sanctions mitigate the problem of dishonesty, even if they do not quite solve it.

In the United States, critics argue that existing sanctions fall short for deterring falsehoods, especially those propagated online. Unlike manufacturers and doctors, the purveyors of online falsehoods are often difficult to find, live outside the jurisdiction, and are judgment-proof.<sup>10</sup> Challengers often lack standing—and even a cause of action—when trying to sue over false stories causing generalized harm.<sup>11</sup> Suing costs time and money, damages are speculative, and collection is uncertain. The First Amendment shields many speakers, liars included, from lawsuits and regulations, allowing actions only for certain categories of falsehoods that cause cognizable harms.<sup>12</sup>

Recognizing the shortfall, a new wave of literature calls for far-reaching reforms. Leading scholars and politicians have considered a dizzying, and sometimes contradictory, array of regulatory fixes. Information labels, fact checks, expert curation, censorship, signal boosts and shadow bans,<sup>13</sup> platform liability, platform immunity, platform transparency, media subsidies, antitrust tools—the list of reforms goes on. Some of these proposals have caught on with legislators. Recently, Texas passed a law meant to fix perceived anti-conservative bias by limiting social media platforms’ ability to curate content.<sup>14</sup> In stark contrast, President Biden expressed frustration with the failure of platforms to curate content amid the Covid-19 pandemic—“They’re killing people.”<sup>15</sup> The Biden campaign called for the repeal of Section 230 of the Communications Decency Act of 1996 (“CDA”), the federal law that provides a legal shield to interactive computer services for most user-generated content. Recently a group of Democrats proposed legislation in this vein.<sup>16</sup>

We make three contributions to the debate. First, we generate a clear, comprehensive, and concise taxonomy of the various proposals. In our framework, solutions to misinformation can be understood as attempts to accomplish one or more of the following three goals: increase the supply of true information, decrease the supply of false information, or improve people’s ability to know the difference. Much of the problem of misinformation, we argue, can be thought of as a ratio: either too many falsehoods or too few truths. This organizational simplicity in an area full of scattershot proposals is important, not least because it exposes contradictions and tensions within reforms and the

---

<sup>10</sup> On the harms from defamatory remarks, see generally DANIELLE CITRON, *HATE CRIMES IN CYBERSPACE* (2014). On the challenges of dealing with anonymous speech, see Lyrrisa Barnett Lidsky, *Silencing John Doe: Defamation and Discourse in Cyberspace*, 49 *DUKE L.J.* 855 (2000).

<sup>11</sup> See *infra* note 185 and accompanying text.

<sup>12</sup> See, e.g., *United States v. Alvarez*, 567 U.S. 709 (2012) (“When content-based speech regulation is in question, however, exacting scrutiny is required.”).

<sup>13</sup> See, e.g., Alstyn, *supra* note 8, at 23 (proposing that platforms that limit the exposure of accounts produce misinformation.)

<sup>14</sup> John Villasenor, *Texas’ New Social Media Law is Blocked For Now, But That’s Not the End of the Story*, *BROOKINGS* (Dec. 14, 2021).

<sup>15</sup> Quinta Jurecic, *The Politics of Section 230 Reform: Learning from FOSTA’s Mistakes*, *BROOKINGS* (Mar. 1, 2022).

<sup>16</sup> Margaret Harding McGill, *E&C leader talks tech reform with Facebook, Google, Twitter*, *AXIOS* (May 24, 2021).

motivations to pass them. These tensions suggest that something might be missing from our understanding of the problem itself.

Our second contribution is to use the tools of economics to complement and generalize the insights of scholars working in this area, who are mostly experts in constitutional and administrative law.<sup>17</sup> We share their diagnosis that the marketplace of ideas is producing undesirable outcomes.<sup>18</sup> We distill their concerns into two distinct sources of market failure: spillovers and information asymmetries.<sup>19</sup> These two distinct failures call for different solutions.

Spillovers arise when one person's speech affects others. Because of spillovers, we have too much low-quality information and not enough high-quality information in circulation. This helps explain why many reforms focus on the "supply side," that is, on the producers of information. Reforms aim, directly or indirectly, to reduce the production of false speech or promote the production of true speech. However, spillovers are not the whole story. The marketplace of ideas also suffers from a distinct failure of information asymmetries. Such asymmetries arise when speakers know more about the veracity of the information than they share than do their listeners. Audiences are left to wonder what information to trust and what to discard. This calls for "demand side" solutions, meaning solutions focused on information consumers, with fact-checking on social media being a prominent example.

Scholars before us have referenced these market failures, but some of their finer implications have gotten lost. We leverage the economic framework to offer a critical evaluation of some dominant proposals. As we show, some proposals run the risk of being ineffective or even counterproductive, exacerbating the problems they seek to solve. We believe that this evaluation contributes an important element to contemporary debates and provides new tools for improved solutions.

Our final contribution is to develop a solution to misinformation. The solution is general; it could apply to accidental errors in publications, deliberate lies, defamatory statements about individuals, and generalized lies that are not legally cognizable. It works on both market failures. On the supply side, it reduces the incentive to produce low-quality information by making it more expensive (and conversely, increases the rewards to producing high-quality

---

<sup>17</sup> See, e.g., Frederick Schauer, *Free Speech, the Search for Truth, and the Problem of Collective Knowledge*, 70 SMU L. REV. 231 (2017); David Pozen, "Truth Drives Out Lies" and Other Misinformation, KNIGHT FIRST AMEND. INST. AT COLUM. UNIV. (Feb. 9, 2022); MINOW, *supra* note 4; HASEN, *supra* note 4; Genevieve Lakier, *The Invention of Low-Value Speech*, 128 HARV. L. REV. 2166 (2015); Evelyn Douek, *Content Moderation as Administration*, 136 HARV. L. REV. (2022).

<sup>18</sup> As Blocher notes, the marketplace analogy remains "the reigning (if somewhat embattled) justification for free speech." Joseph Blocher, *Institutions in the Marketplace of Ideas*, 57 DUKE L.J. 821, 847 (2008).

<sup>19</sup> While we borrow ideas from economic theory, we do not make any strong assumptions of rationality or perfect competition. See, e.g., *id.*, at 833 (noting the limits of the marketplace of ideas given, among others, "participants' imperfect ability to reason"). To the contrary, we are sensitive to the many cognitive and epistemic problems that prevent individuals from engaging in optimal decision-making. At the same time, we take seriously Professor Lidsky's admonition that to live in a democracy requires some degree of respect and trust in the faculties of ordinary people. See generally Lyriisa B. Lidsky, *Nobody's Fools: The Rational Audience as First Amendment Ideal*, 2010 U. ILL. L. REV. 799 (2010). We emphasize the role of a public that earnestly tries to learn about the world but is often stymied. See *infra* note 187 and accompanying text.

information). On the demand side, the solution aims to overcome information asymmetries by helping people distinguish truths from lies. Unlike fact-checking, our proposal labels communications before they circulate, punishes liars, rewards truth-tellers, and reaches more forms of communication. It does not require constitutional amendments, regulations, or other government involvement. It just requires a clear and complete understanding of why the marketplace of ideas has failed.

The key insight is straightforward: to fight misinformation, speakers must have skin in the game. They must lose something—they must be punished—when they lie. Law often attempts to punish communicators of false information, but it does so bluntly and often ineffectively. Suing and prosecuting individuals is often unsuccessful due to jurisdictional challenges, legal standards, and defendants with few assets to recover; the deep pockets—social media companies and other online hubs—enjoy broad immunity from liability thanks to Section 230. The trick to our mechanism is to have communicators punish *themselves*, and do so voluntarily and frequently. This may sound counterintuitive, but as we emphasize throughout, most speakers do not simply want to produce information—they want listeners to *believe* it. For listeners to believe information, speakers must be credible. One way to gain credibility is to punish oneself for lying.

We dub our mechanism “truth bounties.” In brief, a communicator—we will focus on an editorial board or a freelance writer, but it could be anyone—would publish a story, advertisement, press release, etc., and simultaneously pledge money (say, \$10,000) to a third party. The story would bear an icon indicating the bounty and its amount. Anyone who believes the story to be false could file a challenge. To discourage frivolity, trolling, and strategic action, the challenger would have to pay a fee to the third party, akin to a filing fee in court. Private arbitrators would resolve the dispute, avoiding entanglements with the government and the First Amendment. If the challenger won, she would get the journalist’s bounty, and the loss of the bounty would be publicized. If the challenger lost, the bounty would remain for others to claim.

Truth bounties are to speech what product warranties are to refrigerators.<sup>20</sup> Truth bounties let communicators put skin in the game. We expect speakers to post bounties for roughly the same reasons that manufacturers offer warranties.<sup>21</sup> Truthful communicators will welcome bounties because they send a clear positive signal: they stand by their work. Truth bounties are a surefire way to gain credibility and its attendant benefits: readers, buyers, voters. Because serious people will not lose their bounties—their news is not fake—the system poses little risk for them. Hoaxers, on the other hand, will shun the risk. Someone will successfully challenge their fake news and win the money. Foreseeing this, hoaxers will not post a bounty.

---

<sup>20</sup> Sanford Grossman, *The Informational Role of Warranties and Private Disclosure About Product Quality*, 24 J.L. & ECON. 461 (1981); George A. Akerlof, *The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488, 499–500 (1970).

<sup>21</sup> Research shows that warranties result in higher purchase intentions and higher perceived quality, consistent with theory. *See, e.g.*, Jens Høgreve & Dwayne D. Gremler, *Twenty Years of Service Guarantee Research: A Synthesis*, 11 J. SERV. RES. 322 (2008).

The public would be a principal beneficiary of this system. The truth bounty icon could appear next to communications the moment they circulate; no need to wait days or weeks for a fact check. The icon would tell consumers which stories have bounties and are therefore credible. Rather than resorting to crude heuristics, such as only watching a single TV channel one trusts, consumers could indulge in a richer information diet. With truth bounties, truth and lies can separate rather than mix. Truth bounties would be open to all—anyone could attach a bounty to their speech, and anyone could challenge it. Thus, the system would sidestep the challenges of borders, standing, and jurisdiction, while democratizing the search for truth.

Truth bounties offer a promising and robust solution to a vexing problem.<sup>22</sup> Besides contextualizing truth bounties in light of competing reform proposals, our objective here is to lay the groundwork for an actual, workable system.

One might worry that truth bounties favor the rich and harm the poor. This is an important concern, but we do not think it is well founded. The system would aid consumers of information by helping them sort what is true from what is false free of charge. Insofar as poorer people have less education and fewer alternative mechanisms for filtering misinformation, truth bounties would be especially beneficial. With respect to the production of information by journalists and others, truth bounties would not necessarily be expensive. Honest producers of information would *not* lose their bounties; they would get the money back after a certain time. By signaling credibility, bounties would allow small players to compete in the marketplace of ideas with established, monied interests such as major broadcasting networks and newspapers. One might worry that rich actors could take advantage of the system by placing a bounty on stories that are false. In the best-case scenario the bounty would help the false story catch on. In the worst-case scenario, someone would challenge the story and collect the bounty, but the rich actor would not mind because she has plenty of money to spare. This could happen, but we do not think it would be likely or common. Even if a wealthy actor were willing to bear the loss of bounties, the system would record and publicize her track record—every story that she bountied, every challenge that she lost, and so on. Everyone would see that her stories lack credibility. We will return to these issues below.<sup>23</sup>

The last and perhaps most radical contribution of our paper is optimism. Reading the literature on fake news and misinformation, one cannot avoid an overwhelming feeling of pessimism. Many scholars who are learned in the liberal tradition, committed to the values of a free society, and acutely aware of the history of government overreach, censorship, and discriminatory distribution of

---

<sup>22</sup> We are not the first to argue that contractual devices (including warranties) can promote honesty, but we are the first to develop truth bounties in detail. See Yonathan Arbel, *Slicing Defamation by Contract*, U. CHI. L. REV. ONLINE 109 (2020) (briefly discussing truth bounties in the context of defamation law). In a short blog post we discovered after writing this Article, Robin Hanson briefly discusses “News Accuracy Bonds.” See Robin Hanson, *News Accuracy Bonds*, OVERCOMING BIAS (Sep. 9, 2018, 12:30 PM), <https://www.overcomingbias.com/2018/09/news-accuracy-bonds.html>. Two scholars explored a mechanism akin to truth bounties for expert witnesses. Robert Cooter & Winand Emons, *Truth-Bonding and Other Truth-Revealing Mechanisms for Courts*, 17 EUR. J. LAW & ECON. 307 (2004). For insightful discussions of fake news, credibility, and incentives, see generally Van Alstyne, *supra* note 8, at 27-31; Daniel Hemel & Ariel Porat, *Free Speech and Cheap Talk*, 11 J. LEG. ANALYSIS 46 (2019).

<sup>23</sup> See *infra* Part IV.B.

access to speech, have resigned themselves to the inevitability of speech suppression. A prominent example is Dean Chemerinsky who confessed his apostasy: “I still believe in the premise of the First Amendment—that more speech is better”, and then added, “But ever more, I realize that it is a matter of faith, and the internet may challenge that faith for all of us.”<sup>24</sup> This pessimism may be premature. Thinking beyond the hands-off/hands-on dichotomy of either laissez-faire policies or centralized regulation could help us imagine new solutions. Truth bounties demonstrate the value of institutional designs that break this binary mold. Truth bounties offer an intermediate position and demonstrate that through the building of institutions and market design we can realize important social goals.

---

<sup>24</sup> Chemerinsky, *supra* note 5, at 15.

## I. MISINFORMATION AND PROPOSALS FOR REFORM

Misinformation is an ancient problem. Plato worried about deception and manipulation in politics over 2,000 years ago.<sup>25</sup> Sellers exaggerate, puff, or outright lie about the quality of their goods and services, and presumably they have done so for centuries. However, misinformation seems especially salient today. Publication and dissemination have never been easier. Perhaps as a consequence, “fake news,” meaning false or misleading information presented as accurate reporting, circulates widely on social media. New technology allows for “deep fake” videos that depict real people saying and doing things they never said or did. Changing markets have weakened traditional journalism and investigative reporting, especially at the local level.<sup>26</sup> The scope of private statements, alongside their permanence on the internet, amplify the reach of defamatory statements. Together these developments make it hard for people to assess the veracity of information.

The stakes are high. False claims about election fraud led to an assault on the U.S. Capitol.<sup>27</sup> False claims about COVID have led people to reject valuable vaccines and ingest alternative medicines of doubtful efficacy.<sup>28</sup> Absurdly false claims about politicians engaged in the sex trafficking of minors caused a gunman to storm a restaurant.<sup>29</sup> False claims about Dominion’s voting machines have led to threats against the lives of the company’s management and weakened trust in the democratic process. False claims about Nazis helped Russia (attempt to) justify its invasion of Ukraine.<sup>30</sup>

Given the stakes, many scholars have proposed urgent reforms. This Part canvasses some of those proposals. We cannot do justice to all of the promising reforms on the table, but we can summarize some of the most common and compelling arguments. Understanding them will clarify and distinguish our approach, which we will develop later.

To organize the various proposals in a common framework, we focus on the root problem: people are exposed to a mix of true and false information and

<sup>25</sup> See generally PLATO, REPUBLIC. See also Donald Lateiner, “Bad News” in Herodotos and Thonkydides: Misinformation, Disinformation, and Propaganda, 9 J. ANCIENT HIST. 53 (2021).

<sup>26</sup> JAMES T. HAMILTON, DEMOCRACY’S DETECTIVES: THE ECONOMICS OF INVESTIGATIVE JOURNALISM 12-33 (2017).

<sup>27</sup> Steve Inskeep, *Timeline: What Trump Told Supporters For Months Before They Attacked*, NPR (Feb. 8, 2021); Atlantic Council’s DFRLab, *#StopTheSteal: Timeline of Social Media and Extremist Activities Leading to 1/6 Insurrection*, JUST SECURITY (Feb. 10, 2021).

<sup>28</sup> Brian Stelter & Virginia Langmaid, *Nearly 80% of Americans Have Been Exposed to Covid Misinfo, and Many Don’t Know What to Believe, Survey Says*, CNN BUSINESS (Nov. 9, 2021); Bryan Sullivan, *Fox News Faces Lawsuit For Calling COVID-19 A ‘Hoax’*, FORBES (Apr. 10, 2020).

<sup>29</sup> Jessica Gresko, *‘Pizzagate’ Gunman in DC Sentenced to 4 Years in Prison*, AP (June 22, 2017).

<sup>30</sup> Daniel Funke, *Fact Check: Putin’s Claims Justifying War In Ukraine Are Baseless, Experts Say*, USA TODAY (Mar. 30, 2022, 2:16 PM), <https://www.usatoday.com/story/news/factcheck/2022/03/30/fact-check-why-putins-claims-justifying-war-ukraine-baseless/7089270001/>; Rachel Treisman, *Putin’s Claim of Fighting Against Ukraine ‘Neo-Nazis’ Distorts History, Scholars Say*, NPR (Mar. 1, 2022) <https://www.npr.org/2022/03/01/1083677765/putin-denazify-ukraine-russia-history>.

cannot distinguish between the two. One way to address the problem is to improve the ratio of true to false information. As the ratio improves, the ability to distinguish truth from falsity becomes less important.<sup>31</sup> To see this clearly, consider the limit case: if all information in circulation is true, people's ability to screen out false information becomes irrelevant. An alternative approach is to help people distinguish between true and false information. Based on this, we divide reform proposals into three categories: increasing the numerator of true information, decreasing the denominator of false information, and assisting people with making the distinction.

### A. Increasing the Supply of True Information

In 2021, Professor Martha Minow, the former Dean of Harvard Law School, published an important book titled *Saving the News*.<sup>32</sup> The book received considerable attention,<sup>33</sup> understandably given its lofty ambitions. The core argument is that “the press” has suffered in recent decades and fallen from its “golden age,” which Minow identifies as the era between 1960 and 1980.<sup>34</sup> Because of digital news and algorithms that tailor content to individuals, many people live in echo chambers, giving them “few opportunities to learn, understand, or believe what others are hearing as news.”<sup>35</sup> These developments have led to a democratic deficit, as “trust in news” is “essential in a democratic society.”<sup>36</sup>

Minow offers many proposals to help traditional news. One proposal is to create a royalty system for news shared online. The goal of this proposal is to compensate news creators, whether The New York Times or the non-profit Reveal, for their efforts when their reports are shared and published on social media.<sup>37</sup> Minow's approach would involve the robust enforcement of “intellectual property rights for news” as this “means providing compensation to producers that would help sustain the reporting and writing of material that otherwise is at risk as conventional journalism falters.”<sup>38</sup> The hope is that these augmented resources would encourage the production of high-quality reporting.

---

<sup>31</sup> We should distinguish between, on the one hand, sorting truths from falsehoods and, on the other hand, the costs from failing to do so accurately. It might be better to believe a hundred small lies than to fall for a single big one. See Yonathan A. Arbel & Murat Mungan, *Defamation Law and Bayesian Audiences*, J. LEG. STUD. (Forthcoming, 2022).

<sup>32</sup> MINOW, *supra* note 4.

<sup>33</sup> See, e.g., Kevin M. Lerner, *The News is Dead, Long Live the News!*, BOS. REV. (Aug. 10, 2021), <https://bostonreview.net/articles/the-news-is-dead-long-live-the-news/>; Alex Dalton, *The Former Harvard Law Dean Who Wants Government to Save the News Business*, WASH. MONTHLY (July 26, 2021), <https://washingtonmonthly.com/2021/07/26/the-former-harvard-law-dean-who-wants-government-to-save-the-news-business/>; *Alabama Law Hosts Top Constitutional Law Professors for First Amendment Roundtable Discussion*, ALA. L. SCH. (Oct. 20, 2021), <https://www.law.ua.edu/blog/news/alabama-law-hosts-top-constitutional-law-professors-for-first-amendment-roundtable-discussion/>.

<sup>34</sup> MINOW, *supra* note 3, at 4.

<sup>35</sup> *Id.*

<sup>36</sup> *Id.* at 7.

<sup>37</sup> *Id.* at 104-106.

<sup>38</sup> *Id.* at 107

This proposal aims to increase the supply of accurate information in circulation. As the supply increases, the probability of a particular communication being true should increase, and fewer people should be duped by misinformation, regardless of whether they are adept at distinguishing reliable from unreliable sources.

Whether Minow's proposal would succeed is uncertain. If platforms have to pay license fees for the sharing of quality content, they might prioritize the sharing of unlicensed content. While it costs to produce quality journalism, QAnon and other providers of misinformation gladly license their merchandise for free.<sup>39</sup> Moreover, if sharing news costs platforms, they might prioritize sharing only those stories likely to generate clicks and ad revenue—revenue that they could use to pay the news creators. This could create a dismal equilibrium that incentivizes the production of sensationalist stories, click-bait headlines, and culture-war materials. Thus, proposals that rely on IP enforcement could backfire by increasing the proportion of false or low-quality stories shared on social media. Paywalled journalism might provide an analogue. By generating revenue, paywalls fund the production of quality journalism, but they dampen its distribution.<sup>40</sup> Minow is aware of this risk. She offers to solve it by retiring IP rights after two years.<sup>41</sup> Most of the value of news, however, comes from their immediate consumption. Expansive IP alone could distort incentives, dampen the spread of news, and encourage problematic journalistic practices.

Whatever the merits of Minow's specific proposal, the basic intuition behind it seems sound: to encourage the production of high-quality journalism, we must direct more resources to it. Others have offered similar proposals. Professors Sunstein and Hasen, for example, have each suggested subsidizing journalism.<sup>42</sup> Professor Leiter supports a revival of the “fairness doctrine,” which would require information providers like broadcast media to give major political parties equal time when addressing public issues.<sup>43</sup> Among other

---

<sup>39</sup> See *Berisha v. Lawson*, 141 S. Ct. 2424, 2427 (2021) (Gorsuch, J., dissenting) (“the distribution of disinformation—which costs almost nothing to generate—has become a profitable business while the economic model that supported reporters, fact-checking, and editorial oversight has deeply eroded”).

<sup>40</sup> See, e.g., Mark Hill, *Paywalls, Newsletters, and the New Echo Chamber*, WIRED (Dec. 7, 2020), <https://www.wired.com/story/paywalls-newsletters-and-the-new-echo-chamber/> (Quoting journalism professor Damian Radcliffe “people who are priced out of news . . . will be pushed towards free news, some of which is more dubious in nature”).

<sup>41</sup> MINOW, *supra* note 4, at 107 (“To mitigate [concerns with paywalls], the right to compensation could expire two years from the date of first publication.”).

<sup>42</sup> HASEN, *supra* note 4, at 28, 153-54; Richard L. Hasen, *How to Keep the Tide of Fake News from Drowning Our Democracy*, N.Y. TIMES (Mar. 7, 2022), <https://www.nytimes.com/2022/03/07/opinion/cheap-speech-fake-news-democracy.html>;

CASS R. SUNSTEIN, *DEMOCRACY AND THE PROBLEM OF FREE SPEECH*, 19-21, at 68-75, 89-91 (1993) (discussing normative and policy arguments that would support a regulatory scheme that subsidizes dissemination of legitimate news and information).

<sup>43</sup> Brian Leiter, *The Epistemology of the Internet and the Regulation of Speech in America*, GEO. J.L. & PUB. POL'Y, (forthcoming 2022) [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3939948](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3939948) (last revised Jan. 10, 2022).

effects, the fairness doctrine could increase the supply of accurate information by allowing for real-time challenges to spurious or unsupported claims.

## B. Decreasing the Supply of False Information

Having described proposals to increase the supply of truthful information, we next consider proposals to decrease the supply of false or misleading information. Stricter defamation laws offer one method. Many scholars support such a change.<sup>44</sup> Justices Thomas and Gorsuch would loosen federal constitutional constraints and give states greater discretion to regulate defamation.<sup>45</sup> Presidents Biden and Trump voiced dissatisfaction with what they believe is too little accountability for speech in social and traditional media.<sup>46</sup> Pundits across the political spectrum have expressed similar arguments.<sup>47</sup>

Stricter defamation laws might be helpful. Expanding the range of cognizable legal harms, reducing evidentiary burdens, allowing lawsuits by public figures, and generally making it easier to recover damages from people who spread lies should discourage lying. But this approach has important limits. Defamation law only penalizes false information that harms the reputations of specific people or entities. Defamation law does not reach false information in general, as with the statement, “the war in Ukraine is fake.”<sup>48</sup> Furthermore, expansive defamation laws come with familiar problems, such as the general suspicion of regulation and the risk from letting government officials influence or even decide what is true or false. Beyond these familiar limitations, we offer

---

<sup>44</sup> For a few examples, see Sunstein, *supra* note 5, at 389 (arguing that “New York Times Co. v. Sullivan . . . looks increasingly anachronistic”); Cristina Tilley, *(Re)categorizing Defamation*, 94 TUL. L. REV. 435 (2020); Glenn Reynolds, *Rethinking Libel for the Twenty-First Century*, 87 TENN. L. REV. 465, 465 (2020) (calling for reform); JUSTIN HENDERSON, *THE TORTS PROCESS*, 856 (9th ed. 2020) (“Recent years have seen growing dissatisfaction with . . . the law of defamation.”); David A. Anderson, *Is Libel Law Worth Reforming?*, 140 U. PA. L. REV. 487, 550 (1991) (“The present law of libel is a failure.”).

<sup>45</sup> *Berisha v. Lawson*, 141 S. Ct. 2424 (2021) (Gorsuch, J., dissenting); *McKee v. Cosby*, 139 S. Ct. 675, 682 (2019) (Thomas, J., concurring).

<sup>46</sup> Rachel Lerman, *Social Media Liability Law is Likely to Be Reviewed under Biden*, WASH. POST (Jan. 18, 2021); Michael M. Grynbaum, *Trump Renews Pledge to ‘Take a Strong Look’ at Libel Laws*, N.Y. TIMES (Jan. 10, 2018), <https://perma.cc/M2XJ-JW8M>; Donald J. Trump (@realDonaldTrump), TWITTER (Sept. 5, 2018, 6:33 AM).

<sup>47</sup> See, e.g., Bruce Fein, *End the First Amendment Sanctuary for Fake News*, THE AM. CONSERVATIVE (Feb. 27, 2019, 1:00 PM), <https://perma.cc/CUL8-LC34>; Paul Schindler, *Hoylman Said Stronger Law Would Protect Lincoln Project’s Ivanka-Jared Billboards*, GAY CITY NEWS (Oct. 29, 2020), <https://perma.cc/KUD9-L9QN>.

<sup>48</sup> Defamation law bars false statements about groups unless “the group or class is small such that the matter can be reasonably understood to refer to [a specific] member”. RESTATEMENT (SECOND) OF TORTS § 564A (AM. LAW INST. 1977). It is common to view the maximal group size as consisting of 25 members. See *O’Brien v. Williamson Daily News*, 735 F. Supp. 218, 223 (E.D. Ky. 1990), *aff’d*, 931 F.2d 893 (6th Cir. 1991). Modern interpretations of the First Amendment seem to bar the possibility of creating liability for group libel, notwithstanding the Supreme Court’s decision in *Beauharnais v. People of State of Ill.*, 343 U.S. 250 (1952). See *Brysk v. Herskovitz*, 142 S. Ct. 1369, 212 L. Ed. 2d 326 (2022) (“While the decision [in *Beauharnais*] has never explicitly been overruled, it appears that the case has been limited to its precise facts in subsequent decisions of the Supreme Court.”).

later a more comprehensive critique. For now, it is enough to recognize that defamation law can only do so much.<sup>49</sup>

Stricter defamation laws focus on the *production* of false information. Different proposals aim to limit the *dissemination* of false information. For consumers, the effect would be the same: a smaller proportion of the information they encounter is false. The mechanism, however, is starkly different.

Limiting dissemination shifts the focus from individual writers and news media to content platforms—mainly, social media platforms.<sup>50</sup> A modern turn in the scholarship looks at platforms as sites of administration and governance of free speech rights, whose decisions are influenced but not always determined, by the First Amendment.<sup>51</sup> Under some proposals, Facebook, for example,<sup>52</sup> could “prioritize[e] authoritative news sources” through algorithms or similar means, and “downrank[ ] . . . deceptive content.”<sup>53</sup> This and other proposals use the term “curation” or “content moderation,” but they do not always mean platform self-governance of its content—i.e., standard content moderation. While some call for more robust, but content-agnostic, internal self-control,<sup>54</sup> others envision curation under government mandates, which can be fairly characterized as the outsourcing of censorship (or, in Jack Balkin’s terminology, “collateral censorship.”).<sup>55</sup> An effective, and arguably successful, example of private corporate power over online speakers came in the wake of January 6<sup>th</sup> when Google, Facebook, and Twitter limited or removed QAnon content from their platforms.<sup>56</sup> One study found a steep decline in internet discussions around QAnon following these actions.<sup>57</sup> At the same time, this curation met some political backlash, most recently in the form of a Texas law that sought to make it illegal.<sup>58</sup>

---

<sup>49</sup> See *infra* Part IV.A.

<sup>50</sup> Professor Balkin calls this “New School speech regulation.” See Jack M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 UC DAVIS L. REV. 1149 (2018).

<sup>51</sup> Evelyn Douek, *Content Moderation as Administration*, 136 HARV. L. REV. 3 (2022); Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2020).

<sup>52</sup> Mark Verstraete, Derek E. Bambauer & Jane R. Bambauer, *Identifying and Countering Fake News*, 73 HASTINGS L.J. 821 (2022).

<sup>53</sup> Pozen, *supra* note 17.

<sup>54</sup> See, e.g., Danielle Citron, *Cyber Civil Rights*, 89 B.U. L. REV. 61 (2009).

<sup>55</sup> Balkin, *supra* note 50, at 1177 (“Collateral censorship in the digital era involves nation states putting pressure on infrastructure providers to censor, silence, block, hinder, delay, or delink the speech of people who use the digital infrastructure to speak.”). Beyond censorship, some people suggest more radical reforms, such as engaging antitrust authorities to regulate platforms. See, e.g., MINOW, *supra* note 4; Amy Kapicziak, *Freedom from the Marketplace of Speech*, KNIGHT FIRST AMENDMENT INST. AT COLUM. UNIV. (Feb. 14, 2022), <https://knightcolumbia.org/blog/freedom-from-the-marketplace-of-speech>; Hasen, *supra* note 4, at 130.

<sup>56</sup> Jared Holt & Max Rizzuto, *QAnon’s Hallmark Catchphrases Evaporating from the Mainstream Internet*, MEDIUM: DFRLAB (May 26, 2021), <https://medium.com/dfrlab/qanons-hallmark-catchphrases-evaporating-from-the-mainstream-internet-ce90b6dc2c55>.

<sup>57</sup> *Id.*

<sup>58</sup> CENSORSHIP OF OR CERTAIN OTHER INTERFERENCE WITH DIGITAL EXPRESSION, INCLUDING EXPRESSION ON SOCIAL MEDIA PLATFORMS OR THROUGH ELECTRONIC MAIL MESSAGES, 2021 Tex.

To encourage curation, lawmakers could increase legal exposure for false information published on platforms. To accomplish this, many scholars support reforming Section 230 of the Communications Decency Act.<sup>59</sup> In brief, Section 230 immunizes platforms from liability for speech disseminated through their systems. Many commentators support eliminating Section 230 protections<sup>60</sup> and imposing tort liability on websites that publish “foreseeably harmful” content.<sup>61</sup> Danielle Citron is a leading voice in this area. In a number of articles she has charted a course for the redrafting of Section 230. In her view, Section 230 performs a vital role, because it allows platforms to moderate content without risking legal exposure, thus encouraging the creation of online communities with distinct character. She, however, forcefully rejects the broad protections afforded to platforms that host illegal content. In her view, platforms should be liable for illegal content if they cannot show that they have taken “reasonable steps to address unlawful uses ... that clearly create serious harm to others.”<sup>62</sup>

Legal scholars Mark Verstraete, Jane Bambauer, and Derek Bambauer have a different view. They call for *expansion* of Section 230 protections on the ground that reducing “legal liability for internet platforms” will “encourage intermediaries to filter fake news without risk of lawsuit or damages.”<sup>63</sup> They suggest a model for platforms that would be run by an “elite staff of editors and journalists ... [who would] make difficult editorial judgments about propaganda.”<sup>64</sup> The model entity they propose for this delicate task is the BBC.<sup>65</sup>

These proposals face headwinds. Without adversarial or at least deep investigative processes, platforms cannot necessarily determine the truth or falsity of stories, especially when the stories have “a kernel of truth that enables their creators to artfully mix fact and fiction in a way that upends traditional

---

Sess. Law Serv. 2nd Called Sess. Ch. 3 (H.B. 20) (VERNON’S); Andrew Zhang, *Texas Law Prohibiting Social Media Companies from Banning Users Over Their Viewpoints Reinstated by Appeals Court*, THE TEXAS TRIBUNE (May 11, 2022), <https://www.texastribune.org/2022/05/11/texas-social-media-law-reinstated/>.

<sup>59</sup> Evelyn Douek, *Governing Online Speech: From “Posts-As-Trumps” to Proportionality and Probability*, 121 COLUM. L. REV. 759, 767 (2021), (“Section 230 of the Communications Decency Act —is increasingly under siege across the political spectrum, with its reform seemingly imminent”).

<sup>60</sup> MINOW, *supra* note 4, at 104-138; Brian Leiter, *Cleaning Cyber-Cesspools: Google and Free Speech*, in THE OFFENSIVE INTERNET: SPEECH, PRIVACY, AND REPUTATION, 161-62 (Martha Nussbaum and Saul Lemore eds., 2010).

<sup>61</sup> Leiter, *supra* note 30, at 25.

<sup>62</sup> See, e.g., Citron, *supra* note 54; Danielle Citron, *How to Fix Section 230*, B.U.L. REV. (forthcoming) (last revised 17 Mar. 2022); Danielle Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity*, FORDHAM L. REV. (2017); Danielle Keats Citron & Mary Anne Franks, *The Internet as Speech Machine and Other Myths Confounding Section 230 Reform*, 2020 U. CHI. LEGAL F. 45, 71 (2020).

<sup>63</sup> Verstraete, Bambauer & Bambauer, *supra* note 52, at 23-25.

<sup>64</sup> *Id.* at 27.

<sup>65</sup> *Id.* at 27; see also Brett Frischmann, *Understanding the Role of the BBC as a Provider of Public Infrastructure*, YESHIVA UNIV. JACOB BURNS INST. FOR ADVANCED LEGAL STUD. (Jan. 11, 2017), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2897777](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2897777) (calling for the BBC to create a social media network). Note that the BBC does not have an unblemished record. Mike Thomson, *A Very British Coup*, BBC (Aug. 22, 2005), [https://www.bbc.co.uk/radio4/history/document/document\\_20050822.shtml](https://www.bbc.co.uk/radio4/history/document/document_20050822.shtml) (“the BBC was used to spearhead Britain’s propaganda campaign.”).

modes of debunking information.”<sup>66</sup> When making hard choices that involve discretion, it will be all but impossible for platforms not to consider their own commercial interests.<sup>67</sup> Censorship and curation require a special degree of certainty because there is usually no transparency or adversarial process.<sup>68</sup> No one outside the editing rooms knows what got deleted. The proposal to use disinterested elites for censorship or curation will not solve this problem. Verstraete, Bambauer, and Bambauer demonstrate, perhaps inadvertently, the challenge of having an elite group make censorship decisions. They use the lab-leak theory of Covid-19 as an illustration of “fake news.” They argue that despite persistent debunking, the theory “retains its grip on a significant share of Americans.”<sup>69</sup> Since they published their work, however, this theory has been “re-bunked,” meaning some experts have considered it anew, argued for its plausibility, and urged political actors to take it seriously.<sup>70</sup> If even careful scholars who study fake news reach uncertain conclusions about issues as important as a global pandemic, we should worry about the capacity of any actor to censor information carefully, consistently, and accurately.

In addition to questions of competence, many other challenges to curation and censorship loom. Pre-screening all information could be prohibitively costly, but selectively screening after dissemination might be useless. By then the information has already circulated. More generally, many people frown on regulation and coercion. Censorship, collateral censorship, and even heavy-handed content moderation run contrary to democratic values and the so-called marketplace of ideas.<sup>71</sup> Public actors can abuse censorship for personal advantage, as when Vladimir Putin silences dissent.<sup>72</sup>

In sum, many proposals aim to decrease the supply of false or misleading information. In theory, this strategy should expose people to more truth. In practice, it faces challenges. This approach requires powerful actors, whether the government, large platforms, or both, to screen information. Some solutions are more moderate, but middle-of-the-road solutions can only screen the most egregious forms of disinformation.

---

<sup>66</sup> Verstraete, Bambauer, & Bambauer, *supra* note 52.

<sup>67</sup> Jack Balkin notes how central curation is to the business model of online platforms. “Social media companies . . . realized that a substantial aspect of their product was creating a hospitable environment for end-users.” Balkin, *supra* note 50, at 1149.

<sup>68</sup> Klonick, *supra* note 51, at 1635-1649.

<sup>69</sup> Leiter, *supra* note 30 at 37.

<sup>70</sup> Amy Maxmen & Smriti Mallapaty, *The COVID Lab-leak Hypothesis: What Scientists Do and Don't Know*, NATURE (June 8, 2021); Alexander Smith, *China Slams New WHO Report Suggesting Further Investigation Into Covid 'Lab Leak' Theory*, NBC NEWS (June 10, 2022, 12:38 PM), <https://www.nbcnews.com/news/world/covid-19-urges-investigation-chinese-wuhan-lab-leak-theory-rcna32910>.

<sup>71</sup> *See* Se. Promotions, Ltd. v. Conrad, 420 U.S. 546, 553(1975)(“Our distaste for censorship—reflecting the natural distaste of a free people—is deep-written in our law.”).

<sup>72</sup> Leiter, *supra* note 13, at 16 (“The primary reason to be skeptical of regulation of speech is the reliability of regulators, who often have bad motives for suppressing speech.”); *see also* MILTON, AREOPAGITICA 745 (“[in the search for truth, we must not] set an oligarchy of twenty engrossers over it, to bring a famine upon our minds again, when we shall know nothing but what is measured to us by their bushel.”).

### C. Distinguishing True and False Information

If people could sort truths from lies, we could worry less about producing more of the former and less of the latter. If people were sufficiently savvy, even a gush of misinformation and propaganda could not lead them far astray. This idea is reflected in reform proposals that aim to make information consumers better at making distinctions.

Much work in this spirit focuses on labeling. Social media platforms like Twitter and Facebook already label some posts as “misleading” or “false.”<sup>73</sup> Recently Twitter began flagging tweets that link to Russian state-sponsored media. Private organizations like Ad Fontes Media and NewsGuard rate the reliability of news sources.<sup>74</sup>

Scholars have pushed for more. Professor Hasen wants mandatory labels on altered videos or audio, if and when the technology for detecting such manipulation becomes available.<sup>75</sup> Professor Wood and Commissioner Ravel want mandatory disclosure of the sources of political speech on the internet.<sup>76</sup> Others want labels on information akin to nutrition facts on packaged food.<sup>77</sup>

Labeling is often lauded for its light-touch approach. The labeler—typically a platform or search engine—indicates that the information is contested, inaccurate, or triggering, sometimes with a link to an authoritative (or seemingly-authoritative) source, such as a government agency. Because no information is deleted or redacted, labeling does not raise the same objections as outright censorship. To clarify the point, consider the harm from type-1 and type-2 errors. With censorship, some truths get silenced, and some falsehoods slip through. Mistakes are inevitable. However, mistakes in labeling seem less harmful than mistakes in censoring.

There is much to like about labels. However, they are not a panacea. According to Professor Pozen, “warning labels, fact checks, corrections, criticisms, and the like . . . have disappointed in countless discrete domains,” and “[w]e shouldn’t expect them to solve a world-historical epistemic crisis.”<sup>78</sup> Pozen is probably right that labeling cannot “solve” the problem of misinformation, but it can help. Some evidence shows that labels are effective,<sup>79</sup> and new and potentially useful innovations in labeling are in development.

---

<sup>73</sup> Rachel Kraus, *Facebook Labeled 180 Million Posts As 'False' Since March. Election Misinformation Spread Anyway*, MASHABLE (Nov. 19, 2020), <https://mashable.com/article/facebook-labels-180-million-posts-false>; Musadiq Bidar, *Twitter Will Label Posts with Misleading Information about COVID-19 Vaccines*, CBS NEWS (Mar. 2, 2021), <https://www.cbsnews.com/news/twitter-covid-19-vaccine-misinformation-labels/>.

<sup>74</sup> AD FONTES MEDIA, INC. <https://adfontesmedia.com/about-ad-fontes-media/> (last visited Mar. 22, 2022).

<sup>75</sup> Hasen, *supra* note 3, at 144.

<sup>76</sup> See generally Wood & Ravel, *supra* note 5.

<sup>77</sup> Matthew Spradling, Jeremy Straub & Jay Strong, *Protection from 'Fake News': The Need for Descriptive Factual Labeling for Online Content*, 13 FUTURE INTERNET 142 (2021).

<sup>78</sup> Pozen, *supra* note 17.

<sup>79</sup> See *infra* notes 158-160 and accompanying text.

## II. FAKE NEWS IN THE MARKETPLACE OF IDEAS

Misinformation is a pressing social problem. But why exactly did this problem emerge? Why do some people *produce* false information and, more challenging, why do other people choose to *believe* or *share* it?<sup>80</sup>

Many analyses today pin the rise of misinformation on the internet, the greater ease of publishing and sharing information, and information's digital permanence. None of these factors, however, goes to the root of the issue. As Bryan Caplan notes, these explanations suffer because they “focus[] exclusively on the flaws of *speakers*, without acknowledging the flaws of the *listeners*.”<sup>81</sup> What we are missing, in other words, is a solid understanding of both the supply *and demand* for false information—an understanding of the whole marketplace of ideas.<sup>82</sup>

The marketplace of ideas metaphor imagines *competition* among speech acts. Just as high-quality products will push inferior alternatives out of the market, high-quality speech will push misinformation out of the market.<sup>83</sup> The idea has a powerful allure, but as noted recently by Rick Hasen, “The marketplace of ideas is experiencing market failure.”<sup>84</sup> Borrowing from economic theory, we argue that the market actually suffers from not one, but *two* distinct failures. It fails once because speech has spillover effects on the broader society, and it fails again because of inherent information asymmetries between speakers and audiences.

Spillovers and information asymmetries are different in their cause, logic, and remedy. Conflating or ignoring these failures is a recipe for misguided and counterproductive reform proposals. We define the market failures and leverage them to illuminate some weaknesses and unintended consequences of common reform proposals. This discussion builds the foundation for our reform, which we believe avoids the pitfalls of some others. We also believe this discussion supplies independent value by providing a framework and some key distinctions to enrich the debate.

---

<sup>80</sup> People may knowingly share false reports for a variety of reasons. Scott Alexander, *The Toxoplasma of Rage*, SLATE STAR CODEX (Dec. 17, 2014), <https://slatestarcodex.com/2014/12/17/the-toxoplasma-of-rage/>.

<sup>81</sup> Bryan Caplan, *Misinformation About Misinformation*, BET ON IT (May 18, 2022), <https://betonit.substack.com/p/misinformation-about-misinformation?s=w>.

<sup>82</sup> Alex Tabarrok makes a similar point: “it’s an equilibrium process. The demand and supply of misinformation both matter.” Alex Tabarrok, *The Demand and Supply of Misinformation*, MARGINAL REVOLUTION (May 20, 2022), <https://marginalrevolution.com/marginalrevolution/2022/05/the-demand-and-supply-of-misinformation.html>.

<sup>83</sup> The metaphor dates to *Abrams v. United States*, 250 U.S. 616 (1919). *See also* RICHARD POSNER, *THE PROBLEMS OF JURISPRUDENCE* 115 (1990) (applauding the “Darwinian test” for ideas).

<sup>84</sup> HASEN, *supra* note 4, at 23. For an earlier statement by an economist, see Ronald H. Coase, *The Market for Goods and the Market for Ideas*, 64 AM. ECON. REV. 384, 385 (Arguing that “there is a good deal of ‘market failure’ in the U.S. marketplace for ideas”).

## A. The Spillover Problem

This section addresses spillovers, a common source of market failure. We begin with a brief overview of spillovers and then discuss the challenges of correcting them. Those challenges are especially acute in the context of speech and information, presenting problems for some suggested reforms.

### 1. On Spillovers

Spillovers arise when people's choices affect others. These outside effects are often neglected by individuals and not given sufficient weight, precisely because the decisionmaker does not bear the full consequences. When a factory pollutes the air, its emissions harm everyone nearby. Because these harms do not affect the factory's bottom line, it may continue operating while imposing this negative spillover. Negative spillovers cause people to engage in more of a harmful activity than they should.<sup>85</sup>

Positive spillovers work similarly. When a homeowner keeps bees or installs a light in a dark alley, she benefits herself and others too.<sup>86</sup> The total benefits of pollinators and light might exceed the costs. However, the homeowner does not enjoy all of the benefits; many of them flow to other people. So, homeowners keep too few bees and install too few lights. Positive spillovers cause too little of the beneficial activity.

Speech suffers from negative and positive spillovers.<sup>87</sup> Misinformation can impose harm on the general public, as when fake reports diminish trust in democratic institutions or social cohesion. High-quality journalism, whistleblowing, inspiring oratory, and other forms of speech have broad benefits. These benefits go beyond the commercial interests of speakers and consumers of their speech. An informed citizenry is a social interest, transcending the private interests of either the daily paper or its readers. So are confidence in the democratic process, institutional legitimacy, the rooting out of corruption, and checks on political excess. The problem is that speakers do not necessarily capture those benefits.<sup>88</sup> Consequently, quality journalism is in short supply. Spillovers distort the market for speech.

### 2. Spillover Critique of Reform Proposals

---

<sup>85</sup> ROBERT D. COOTER & MICHAEL D. GILBERT, PUBLIC LAW AND ECONOMICS 32-35 (2022).

<sup>86</sup> Steven N.S. Cheung, *The Fable of the Bees: An Economic Investigation*, 16 J.L. & ECON. 11 (1973) (investigating market reactions to pollination externalities to apple growers).

<sup>87</sup> See, e.g., Daniel A. Farber, *Free Speech Without Romance: Public Choice and the First Amendment*, 105 HARV. L. REV. 554 (1991).

<sup>88</sup> On the economics of investigative journalism, see Oren Bar-Gill & Assaf Hamdani, *Optimal Liability for Libel*, 2 J. OF ECON. ANALYSIS & POL'Y 1 (2003); Nuno Garoupa, *Dishonesty and Libel Law: The Economics of the "Chilling" Effect*, 155 J. INST'L & THEORETICAL ECON. 284 (1999); Nuno Garoupa, *The Economics of Political Dishonesty and Defamation*, 19 INT'L REV. L. ECON. 167 (1999).

Spillovers diminish the capacity of an unregulated market to produce good results. Regulation and subsidies can help by causing speakers to “internalize” these spillovers and thus improve outcomes.<sup>89</sup> So far, we are in agreement with others. What has not been widely recognized in the speech debate, however, is how easily reforms meant to fix spillovers can make matters worse. Correcting spillovers requires *precise* interventions, and precision poses a difficult, practical problem.

Consider an example. A ranch fouls a nearby stream. If the ranch operates, the owner earns a profit of 10. The farmers downstream pay a cost from the water pollution equal to 14. If the ranch does not operate, no one gains or loses anything. Given these options, the ranch should not operate. The net payoff from not operating equals 0, which exceeds the net payoff from operating ( $10 - 14 = -4$ ). However, absent regulation, the ranch *will* operate.<sup>90</sup> The root problem is, quite literally, a spillover. The rancher does not pay the 14 in costs. They spill over to the neighbors. For the rancher, operating leads to a profit of 10, so she operates, even though the net payoff for society is  $-4$ .

In this example, the market for ranching functions poorly. To correct it, law should impose liability (in tort, or perhaps a tax) on the rancher equal to the social harm she causes,  $-14$ . The liability causes the rancher’s personal calculation (10 profit,  $-14$  from the tax, for a net payoff of  $-4$ ) to match society’s calculation (10 profit,  $-14$  from pollution, for a net payoff of  $-4$ ). The tax induces the rancher to consider all costs and benefits, not just her own, when deciding how to act.

For this strategy to work, liability needs to have a degree of precision. Errors in setting liability can lead to *worse* results. To illustrate, suppose the ranch causes 14 in harm to the neighbors, but the court is expected to impose liability of only 6. For the rancher, operating leads to an expected profit of  $10 - 6 = 4$ , which is better than not operating and earning 0. The rancher operates, even though the net payoff from doing so equals  $-4$ . To generalize, setting liability too low “under corrects,” failing to stop the harmful activity. The opposite problem can arise too. Changing our example, suppose the ranch causes only 6 in harm to the neighbors, but the court erroneously imposes liability of 14. The rancher’s payoff from operating equals  $10 - 14 = -4$ , so she does not operate. But society’s payoff from operating equals  $10 - 6 = 4$ , so she should operate.<sup>91</sup> Setting liability too high “over corrects,” stopping a beneficial activity.

The logic works the same with positive spillovers—only that now, a subsidy rather than a penalty may be required. To illustrate, imagine flower farms. The farms earn profits for the owners and, by supporting pollinators, benefit other

<sup>89</sup> The idea of using taxes to internalize externalities dates to ARTHUR PIGOU, *ECONOMICS OF WELFARE*, 159-175 (1932).

<sup>90</sup> Assuming the transaction costs of bargaining between the rancher and the neighbors downstream are high. If the transaction costs are zero, the efficient outcome will prevail. See Ronald Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1 (1960).

<sup>91</sup> Perhaps she should operate and also compensate the neighbors for their harm.

growers nearby. The optimal subsidy for flower farms equals the size of their positive spillover. If the subsidy is too low, the farms will not operate, even though they should. If the subsidy is too large, some farms will operate when they should not.

This simple analysis illuminates some proposals to address fake news. Earlier we described a proposal to subsidize quality journalism. This proposal works in theory. Like the flower farm in our example, quality journalism creates benefits that flow not just to the journalists and paying readers but to society at large. This positive spillover implies that the unregulated market will produce too little quality journalism. Law can correct this spillover with a subsidy. But what's the proper amount? If we make the subsidy too small, we might mitigate the problem, but we will not solve it. The market will still produce too little quality journalism. If we set the subsidy too high, we create a different, and perhaps less obvious, problem. A too-high subsidy will generate journalism that is socially wasteful and potentially harmful. Think of deep investigations of esoteric issues, fierce and resource-intensive competitions among journalists to scoop each other, and reports that simply check the box of whatever qualifies for a subsidy or the grant. The problem is compounded by the difficulty of setting criteria for these subsidies. The state of Iowa, for example, discovered that 80% of its \$32 million tax credits granted to support movie productions were misspent.<sup>92</sup> Misallocated subsidies could end up supporting partisan efforts, possibly producing more misinformation.

Instead of promoting good information, some of the proposals we canvassed aim to deter bad information. Consider reforming defamation law. Making it easier for victims of defamation to sue and recover damages should discourage lying. The argument works in theory but not necessarily in practice.<sup>93</sup>

Suppose a defamer makes a false statement that harms a person's reputation. The lie generates a benefit for the defamer (financial, psychological) worth 10 and imposes a cost on the victim of 12. When defamation is very hard to prove, the cost of 12 becomes a negative spillover. The defamer gets a benefit from lying and pays no cost, so he keeps lying. Making defamation easier to prove does not necessarily help. If the defamer must pay 12 in damages, then the negative spillover disappears. But suppose the court errs and awards damages below 10. In that case, the defamer will still lie. Or suppose the court awards damages greater than 12. When damages get too high, victims (or people who claim to be victims) can obtain counsel, and speakers clam up. As the risk

---

<sup>92</sup> Office of Auditor of State, Report on Special Investigation of the Film, Television and Video Project Production Program (2010). Available online at <http://publications.iowa.gov/9937/1/1060-2690-0E00.pdf>

<sup>93</sup> In practice, powerful parties have an advantage in using these mechanisms. Anti-SLAPP laws, enacted in 32 states, reflect the recognition that defamation law is routinely abused. However, Anti-SLAPP legislation only offers limited protection from abuse of process. For the sake of argument, we will set these problems aside. Even so, expanding defamation law would not be a panacea.

of errors and high damages grows, journalists might stop reporting.<sup>94</sup> Excessive liability, even if designed to provide redress for innocent victims, can threaten journalism.<sup>95</sup>

Thinking in terms of spillovers highlights another shortcoming of defamation law. Defamation law assumes that the victim suffers all of the harm, but this is too simple. False statements not only harm the victim; they harm the public.<sup>96</sup> To illustrate, if a defamatory statement causes a business to lose 20% of its profits, this harms the business. But it must also harm some customers who were misled about the business and took their money elsewhere. This “elsewhere” may be inferior; after all, customers could have transacted with the alternative business in the first instance but chose not to. The aggregate loss to customers may exceed that of the business. Even when the target of a defamatory statement recovers for her full harm, defamation law is under-compensatory. Negative spillovers persist.

Finally, recall that defamation law only applies to false statements that harm a person’s or entity’s reputation. It does not apply to false statements in general, such as spurious accusations about stolen elections and bad vaccines.<sup>97</sup> Defamation cannot reach and therefore cannot correct negative spillovers associated with such speech.<sup>98</sup>

To summarize, high-quality speech often has positive spillovers, so the free market produces too little of it. Low-quality speech often has negative spillovers, so the market produces too much. By using the tools of economics, we can recast many proposals to address fake news as efforts to correct those spillovers—to make speakers “internalize” more of the benefits and costs associated with their speech. The analysis of spillovers organizes and simplifies much of the debate about fake news. It also reveals shortcomings in some proposals. To correct spillovers, we must price them accurately. Accuracy is difficult enough when the spillover is water pollution that damages crops. It gets much harder when the spillover is information that affects the choices of an unknown number of people in unknown ways. And it is nearing impossible when the benefits are as diffuse and ethereal as trust in democracy, checks and balances, and rule of law.

---

<sup>94</sup> England historically had strict laws and has become known for “libel tourism.” Journalists around the world report receiving extensive legal threats that originate in England. See THE FOREIGN POL’Y CTR., UNSAFE FOR SCRUTINY (Susan Coughtrie & Poppy Ogier eds., 2020). The problem became so severe that Congress unanimously enacted the SPEECH Act, which makes foreign defamation judgments unenforceable if they fail to meet U.S. standards or if they would exceed the bounds of Section 230.

<sup>95</sup> *New York Times Co. v. Sullivan*, 376 U.S. 254, 273 (1964) (“Whether or not a newspaper can survive a succession of such judgments, the pall of fear and timidity imposed upon those who would give voice to public criticism is an atmosphere in which the First Amendment freedoms cannot survive.”).

<sup>96</sup> See generally Yonathan Arbel, *A Reputation Theory of Defamation Law* (Manuscript, on file with author).

<sup>97</sup> See *supra* notes 47-49 and accompanying text.

<sup>98</sup> The lack of regulation, according to some recent work, also means that the harm would be mitigated. See Hemel & Porat, *supra* note 22; Yonathan A. Arbel & Murat Mungan, *The Case Against Expanding Defamation Law*, 71 ALA. L. REV. 453 (2019); Arbel & Mungan, *supra* note 31.

## B. Information Asymmetry

This section addresses information asymmetries, a source of market failure distinct from spillovers. We first describe information asymmetries in general and in the context of speech in particular. Then we focus on a concept central to information asymmetries, credibility. Finally, we use our discussion of credibility to highlight shortcomings in some proposed reforms.

### 1. On (Mis)Information Asymmetry

Information asymmetry arises when one party to an exchange has information that is not available to the counterparty.<sup>99</sup> To illustrate, suppose the seller sells a used car. The buyer wants to buy it but is aware that some used cars are defective. When the parties negotiate, only the seller knows whether her specific car is in good or bad condition. The problem for the buyer is that the seller might not disclose the condition of the car, and even if she does, the buyer cannot take her at her word. After all, the seller might be lying when she says that the car is in good condition. If the buyer can't independently verify the condition of the car, then a mutually-beneficial exchange can fall through.

The same problem applies in the context of speakers and listeners.<sup>100</sup> A speaker—a journalist, advertiser, politician—makes a statement. Some of these statements reveal valuable information known to her personally, like what happened in a private meeting “on the bottom level of an underground garage just over the Key Bridge in Rosslyn.”<sup>101</sup> The speaker speaks presumably because she wants to spread her message. The problem for the speaker is that some other speakers are disingenuous. Alongside quality journalism, there is reporting based on lies, propaganda, and sensationalism. Just as buyers cannot tell which sellers are trustworthy, listeners cannot tell which speakers are trustworthy. In such instances, listeners are increasingly reliant on trust in the editorial board or outlet. But in an age where trust in media outlets is low, such credibility signals are unavailing. This is the information asymmetry; speakers know (or should know) whether their speech is accurate, but listeners do not. Because of information asymmetries, high-quality journalists find it difficult to distinguish themselves from low-quality propagandists.

At bottom, information asymmetry presents a credibility problem—listeners cannot tell which sources to trust, which to discount, and how much. The problem is two sided, afflicting audiences who seek reliable sources and

---

<sup>99</sup> Akerlof, *supra* note 21. See also COOTER & GILBERT, *supra* note 85, at 36-39.

<sup>100</sup> Drawing on these ideas, Rick Hasen suggests that a central threat today is “cheap speech,” low-quality information that is cheap to produce and circulate. In his view, the consequence of cheap speech is the erosion and possible displacement of higher value speech. HASEN, *supra* note 4, at 30-46.

<sup>101</sup> Bob Woodward, *How Mark Felt Became ‘Deep Threat’*, WASH. POST (June 2, 2005), (recounting the Watergate affair).

speakers who want to distinguish their truthful speech from misinformation. In a sense, the credibility problem is more central than spillovers. Imagine a world with speakers who lie and speakers who tell the truth, and suppose the audience knows exactly whom to trust. In this world, no one would believe speakers who share lies, so lies could not damage reputations or otherwise cause harm. Negative spillovers would not exist.

Scholars and others who address misinformation tend to neglect credibility effects, perhaps motivated by the belief that individuals are not discerning consumers of information. Public audiences naively believe what they hear. This view neglects a robust body of research ranging from epistemology and decision theory to evolutionary psychology and child development, from information economics to the sociology of knowledge, and from marketing to folk wisdom. Without going into detail, this body of research shows humans seek credible sources, possess sophisticated cognitive capability to distinguish credible and non-credible sources, and dismiss and discount unreliable speakers.

To give a flavor of these ideas, consider a study on child development. Three-year-old and four-year-old children were exposed to two speakers. Both speakers stated the names of objects that the children could see. However, occasionally they would slip, calling a shovel a towel and a ball a cookie. The researchers found that children not only “wrote off” unreliable speakers, they engaged in more nuanced judgments. Four-year-olds were “able to differentiate between an informant who was 75% accurate and an informant who was 25% accurate and preferred to seek information from the more accurate informant.”<sup>102</sup> This suggests that even young children can keep a mental account of speaker reliability and assign greater credence to sources more likely to produce accurate statements.

Even in the animal kingdom credibility matters. Peahens prefer fit mates whose offspring will survive in the jungle. Consequently, peacocks clamor to advertise their fitness. But talk (really, squawk) is cheap, and peahens cannot know whom to trust. Thus, a signal evolved for reliably sorting competent peacocks from hopefuls: colorful and weighty plumage. Only the fittest of peacocks can survive to sexual maturity with such luggage on their back.<sup>103</sup> Plumage is a credible signal of fitness precisely because it attracts predators and inhibits food gathering.<sup>104</sup> The principle relates directly to misinformation. Listeners, even peahens, look for credible signals. Only those who send credible (and costly) signals, like bright and heavy feathers, can be believed.<sup>105</sup>

---

<sup>102</sup> Elizabeth S. Pasquini et al., *Preschoolers Monitor Accuracy*, 43 DEVELOPMENTAL PSYCH. 1216, 1223 (2007).

<sup>103</sup> Dustin J. Penn & Szabolcs Számadó, *The Handicap Principle: How an Erroneous Hypothesis Became a Scientific Principle*, 95 BIOLOGICAL REV. 267 (2020).

<sup>104</sup> This is the “Handicap Principle,” a widely accepted theory in evolutionary biology. A. ZAHAVI & A. ZAHAVI, *THE HANDICAP PRINCIPLE: A MISSING PIECE OF DARWIN’S PUZZLE*, 229 (1997) (“The investment — the waste itself — is just what makes the advertisement reliable.”). For other examples, consider an elk’s weighty antlers or a gazelle’s instinct to jump straight up upon seeing a predator.

<sup>105</sup> See A. Grafen, *Biological Signals as Handicaps*, 144 J. THEORETICAL BIOLOGY, 517 (1990).

Back to humans, we find evidence that individuals invest intensively in credible cues. For example, in labor markets, employers search for competent employees. But some employees misrepresent their competence, making the search difficult. In one famous model, job seekers use their level of education to advertise their competence.<sup>106</sup> Unlike standard models where better education implies higher skills, this model uses education to advertise innate ability. To simplify, if good workers find school less taxing and bothersome than bad workers, then (under certain assumptions) bad workers choose less schooling than good workers. Observing this, employers would rather hire better-educated workers, even if schooling imparts no job-related skills.<sup>107</sup> For our purposes, the point is that employers *search* for signals that are credible (“I spent years in school”) and discount empty statements (“I’m a good worker”).

In the context of news, we find that individuals are sensitive to credibility cues in nuanced ways. One study shows that a liberal message on a conservative news site commands more credibility than a liberal message on a liberal news site (and the same holds in reverse).<sup>108</sup>

None of these studies or others purport to show that people are perfect at credibility judgments. We have all had a five-minute conversation with the average voter, and we all make mistakes.<sup>109</sup> People are not always motivated to search for the truth.<sup>110</sup> A large scale study measured attitudes towards immigration among thousands of participants.<sup>111</sup> People often exaggerate the number of immigrants in the country. The study measured attitudes towards the desirable scope of immigration before and after informing respondents on the actual number of immigrants. Respondents indeed changed their mind about the number of immigrants, but not on their attitudes towards immigrants. Correcting people’s misperceptions with accurate information had no effect.<sup>112</sup> Still, these local failures should not distract from the remarkable ability of humans to make complex credibility judgments in many circumstances. We process many types of credibility cues, often unconsciously and rapidly, and reject or discount statements made by unreliable sources.<sup>113</sup> Critically, people are sensitive to speaker incentives, recognizing that costly signals are more reliable

<sup>106</sup> Michael Spence, *Job Market Signaling*, 87 Q. J. ECON. 355 (1973).

<sup>107</sup> For supporting evidence, see BRYAN CAPLAN, THE CASE AGAINST EDUCATION, 96-123 (2018).

<sup>108</sup> Megan Duncan, *What’s in a Label? Negative Credibility Labels in Partisan News*, JOURNALISM & MASS COMMUN. Q. (Oct. 13, 2020), <https://journals.sagepub.com/doi/10.1177/1077699020961856>.

<sup>109</sup> This statement is attributed to Winston Churchill. Michael Richards, *Red Herrings: Famous Quotes Churchill Never Said*, INT’L CHURCHILL SOC. (June 9, 2013), <https://winstonchurchill.org/publications/finest-hour/finest-hour-141/red-herrings-famous-quotes-churchill-never-said/>.

<sup>110</sup> See, e.g., Nyhan, *supra* note 3, at 226 (“many seem especially susceptible to misperceptions that are consistent with their beliefs, attitudes, or group identity”).

<sup>111</sup> Daniel J. Hopkins, John Sides, & Jack Citrin, *The Muted Consequences of Correct Information about Immigration*, 81 J. POL. 315–320 (2019).

<sup>112</sup> *Id.*

<sup>113</sup> See also Nyhan, *supra* note 3, at 229 (providing evidence that on many issues individuals learn and converge on true beliefs, even if the process is slow and incomplete).

than cheap ones. We thus think it is a mistake for scholars to neglect credibility issues when they design reform proposals.

## 2. Information Asymmetry Critique of Reform Proposals

Information asymmetries can be hard to overcome. Take the example of the seller and the buyer. The seller's car is high quality, but the buyer thinks otherwise. Can the seller correct this asymmetry simply by stating, promising, or asserting that her car is high quality? Probably not. Talk is cheap, and the seller might lie. In the news context, the same problem applies. It is not enough for a journal to proclaim that it does "honest reporting" or for a pundit to claim that they "tell things as they are." Speakers can swear fidelity to the truth until their faces turn blue. To change minds and behaviors, listeners must believe them. Speakers must be *credible*.

Recognizing the importance of credibility draws attention to the importance of listeners. We must be attentive to listeners' credibility judgments when considering policy. Consider labeling, as when Twitter flags a dubious story. Effective labeling requires (among other things) that the labeler know the truth of the matter. But labelers do not always know the truth, and most people surely recognize this. Moreover, labeler bias looms large. A recent study found that black, transgender, and conservative individuals are targeted most often for content moderation.<sup>114</sup> Consequently, labeling is a double-edged sword. In attempting to address one credibility problem (should readers trust the story?), it introduces a second credibility problem (should readers trust the labeler?).

Even if labelers were trustworthy, and perceived as such, labeling would still have a potential weakness, as the following study shows. Researchers measured the effect of labeling misinformation on the beliefs of 5,271 participants.<sup>115</sup> Consistent with expectations, they found that a negative label made subjects less likely to believe the story.<sup>116</sup> Encouragingly, they found that labels had a strong effect on people whose political view aligned with the story. That is, people were willing to discount a story that supported "their side" if it was flagged.<sup>117</sup> The unexpected finding concerns the effect on *unlabeled* stories.<sup>118</sup> Participants were more likely to believe unlabeled stories than before.

---

<sup>114</sup> Oliver L. Himson et al., *Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas*, 5 PROCEEDINGS OF THE ACM ON HUMAN-COMPUTER INTERACTION 1 (2021). This survey-based analysis cannot distinguish between algorithmic and human curation.

<sup>115</sup> Gordon Pennycook et al., *The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings*, 66 MGMT. SCI., 4944 (2020).

<sup>116</sup> This is consistent with other findings. See Timo Koch et. al., *The Effects of Warning Labels and Social Endorsement Cues on Credibility Perceptions of and Engagement Intentions with Fake News* 21, (June 13, 2021), <https://psyarxiv.com/fw3zq/>; Megan Duncan, *What's in a Label? Negative Credibility Labels in Partisan News*, JOURNALISM & MASS COMMUN Q. (Oct. 13, 2020), <https://journals.sagepub.com/doi/10.1177/1077699020961856>.

<sup>117</sup> Pennycook et al., *supra* note 114. See also Duncan, *supra* note 116, at 49.

<sup>118</sup> The researchers verify their findings in a separate experiment where some stories are labeled as

Apparently, participants conflated stories that had been reviewed and deemed accurate with stories that had never been reviewed. If this finding is robust, then labeling might, all things considered, *worsen* the information environment by making readers too accepting of unexamined stories.

One way to address this problem would be to increase the reach of labeling efforts. If *everything* gets reviewed, then the absence of a label must mean that the information is accurate. This is impossible under existing practices and reform proposals. Worldwide, people and organizations (and bots) produce and share a massive amount of information daily. This information is produced in a decentralized way and shared across multiple platforms and networks. No existing system can label such flows of information comprehensively and in real time. At best, people can label a small subset of the information, often after the fact. To demonstrate, consider the Washington Post's award-winning "Fact Checker." The Fact Checker is a small team of journalists that runs a handful of stories every week investigating statements by important figures.<sup>119</sup> Meanwhile, the Washington Post itself produces over 1,000 stories, videos, and graphics per day.<sup>120</sup>

Algorithms have been proposed as a solution to the scaling problem. Algorithms have much promise but also raise many problems.<sup>121</sup> One is that algorithms are often a black-box that outsiders struggle to understand and that produce biased results. Another is the difficulty of algorithms in making judgments based on a broader context.<sup>122</sup> Once again, labeling attempts to solve one credibility problem by introducing another one, this time about algorithms and their biases.

Moving away from labeling, we next consider censorship, collateral censorship, and content moderation. Putting aside the moral, political, and institutional concerns, the removal or filtering of information can have unintended consequences in the presence of information asymmetries. Like labeling, truth-based content moderation requires that human moderators (or algorithms) have privileged access to the truth. Even in the domains where this is plausible (e.g., issues on which there is a scientific consensus), censorship and moderation can project an aura of reliability on *all* published communications,

---

stories that have not been vetted. They find that subjects place less faith in these unverified stories. Pennycook et al., *supra* note 114, at 4944.

<sup>119</sup> Glenn Kessler, *About the Fact Checker*, WASH. POST (Jan. 1, 2017), <https://www.washingtonpost.com/politics/2019/01/07/about-fact-checker/>.

<sup>120</sup> Robinson Meyer, *How Many Stories Do Newspapers Publish Per Day?*, ATLANTIC (May 26, 2016) <https://www.theatlantic.com/technology/archive/2016/05/how-many-stories-do-newspapers-publish-per-day/483845/>.

<sup>121</sup> See, e.g., Van Alstyne, *supra* note 8, at 7 (Noting that algorithms are subject to an arms race, as producers of false information are likely to find ways to circumvent the filters).

<sup>122</sup> See CITRON, *supra* note 10, at 232 (2014); Danielle Keats Citron, *Section 230's Challenge to Civil Rights and Civil Liberties*, KNIGHT FIRST AMENDMENT INST.n.41 (2018). Algorithms can increasingly approximate humans in reading and assessing data. See Arbel & Becher, *Smart Readers*, GEO. WASH. L. REV. (forthcoming, 2022).

the theory being that misinformation has been screened out.<sup>123</sup> This is not illogical; enhanced trust is a reasonable response to information gatekeeping. But gatekeeping is necessarily imperfect. Some misinformation will slip by the censors, and when it does people might place greater trust in it than ever. Of course, consumers might react differently. They might question the reliability of the censors, in which case they might place *less* trust in filtered content than in unfiltered content. (Consider the reaction of some Russians to state media.)

Whereas the private removal, blocking, or filtering of online speech aims to reduce the supply of false information, subsidies aim to increase the supply of truthful information. If subsidies work, they lead to more accurate information in circulation and enhanced trust. But enhanced trust means that the occasional false story will penetrate more deeply. In such cases, misinformation becomes *more* persuasive than before, even if it constitutes a smaller share of all information. (Censorship raises the same problem.) For consumers who are not passive, who attempt to make credibility determinations, and who know something about the law—for example, that journalists get subsidies—the effect of subsidies is even harder to predict. Savvy consumers might wonder about the motivations of the actors handing out subsidies. Would conservative readers trust newspapers that receive subsidies from liberal legislators?

Of the common reform proposals, expanding defamation law comes closest to addressing information asymmetries. To see why, consider a recent study.<sup>124</sup> Participants were given a few reports, a newspaper clip, a television screenshot, and a social media post. All of those sources of information related to an article, which the participants also received. The participants were asked to evaluate the credibility of the article. The participants were split into two groups, one instructed that they live in a state where defamation law is effective (liars pay damages), and the other in a state with ineffective defamation law. The study found that defamation laws elicit a clear response. Effective defamation law made participants more trusting of the news, whereas ineffective defamation law led participants to express suspicion and doubt.

This study suggests that defamation law can facilitate credibility judgments by listeners. Knowing that someone was *not* sued for defamation suggests that their speech was truthful. More generally, this study suggests that law can have persuasive power. Changing how we regulate information changes public perceptions of its credibility.

Still, defamation law has shortcomings. One involves its limited reach. Defamation law is inapplicable to general speech or many matters of broad public interest. It requires an identifiable victim who has suffered reputational harm.

---

<sup>123</sup> Arbel & Mungan, *supra* note 98; Hemel & Porat, *supra* note 22.

<sup>124</sup> Yonathan A. Arbel, *The Credibility Effect* (Manuscript on file with authors).

Separate from this problem, defamation law can only do so much to enhance credibility. Like a 1960s television, defamation law produces a low-quality picture. The problem grows from the complexity of the law<sup>125</sup> and, relatedly, the challenge of predicting liability.<sup>126</sup> Suppose the local newspaper runs a story about a teacher alleged to have defrauded the school system. The higher the expected sanction to the newspaper for defamation, the more an observant reader would find the story believable. But can a reader know the expected sanction? Damages in tort law depend on a complex, protracted process, which—in the context of defamation—involves “presumed” damages with little actual proof. Even lawyers struggle to agree on an acceptable range. As a result, liability will be determined by a semi-random assortment of factors concerning the particular effect the publication had on the specific teacher, whether he lost his employment, whether he was retained by a different school, whether he had to undergo therapy, and whether his partner deserted him over the allegation. All of this is entirely opaque to the reader at the time of reading, introducing considerable noise into the signal.<sup>127</sup>

Defamation law can produce a noisy signal in another way. Suppose a court correctly finds for the victim in a defamation suit. On the optimistic view, this makes the community adopt a more favorable view of the victim, perhaps even completely restoring her tarnished reputation. But what if the victim never sues? Victims are heterogenous. Some are rich, sophisticated, powerful, and litigious; others are none of those things. A study presented participants with a report accusing a blogger of trying to blackmail the mayor.<sup>128</sup> One group of participants was told that defamation laws are highly effective, and the other group was told the opposite. Participants were asked to evaluate the credibility of the story assuming the blogger (the victim) did *not* file a lawsuit. The study showed that failing to bring suit acted as a powerful signal. Participants in the effective defamation group were more than three times as likely to believe the accusations against the blogger. Because we know that not all victims sue, even when they have a valid claim, this finding reveals a problem. Failure to sue for defamation, regardless of the reason, may cause people to infer that a false accusation was true.

The last concern marries information asymmetries and spillover concerns. Defamation law can encourage all-or-nothing decisions. Either the publisher decides to publish or not. This is not ideal. In the real world, there is nothing like perfect knowledge. The seller of a car will not always know its mechanical condition with perfect accuracy. A journalist will often have a nuanced understanding of the veracity of a story. Almost no story is unassailably true, and almost no source is beyond reproach. With defamation law, journalists drop stories when there is sufficient doubt—although some of the doubtful stories *do*

---

<sup>125</sup> PROSSER ET AL., PROSSER AND KEETON ON THE LAW OF TORTS, 771 (1988) (“the law of defamation is full of absurdities and anomalies for which no legal writer has ever had a kind word.”).

<sup>126</sup> This issue afflicts audiences *and* speakers, but we focus only on audiences.

<sup>127</sup> Newspapers also face this uncertainty, and it carries chilling consequences for publishing decisions.

<sup>128</sup> Arbel, *supra* note 124.

have merit and *should* be published. But the publisher has to worry about its own liability rather than the public value of the story. And so publishers may shy away from contested issues, their public import notwithstanding.

Ideally, information providers would be able to communicate not just the information itself, but also their degree of confidence. If a reporter knows the source has a conflict of interest, it will be better to communicate this fact alongside the story—even if the reporter still finds the source credible. And if a story is marginal, it might be better to publish it along with the reasons that the editorial board found it doubtful than not to publish it at all. Defamation law does not encourage this type of behavior.

\*\*\*\*\*

In sum, we have argued that speech suffers from two market failures, spillovers and information asymmetries. Because of spillovers, we have too much bad information and not enough good information in circulation. Because of information asymmetries, consumers cannot tell which sources to trust. If consumers do not trust them, high-quality speakers like professional journalists will communicate less or stop communicating entirely. These two market failures call for distinct solutions, each with their own challenges. Fixing spillovers requires a degree of precision in the regulatory response, which in turn requires a hard-to-evaluate assessment of the scope of incentive misalignment. Fixing information asymmetries requires solutions that will increase truth-telling and also public *trust*.

Ideally, solutions would address both of these problems at the same time. We cannot achieve this ideal, but we can move in this direction. We need a strategy for reform that will help people identify trustworthy sources. The strategy must comply with the law, including the First Amendment. It should operate *ex ante*, meaning before information circulates, not after the fact. It must operate at scale, meaning a significant portion of the information in circulation, not just bits and pieces, can be labeled, tagged, or otherwise sorted so that consumers know what to trust. Finally, the strategy must account for credibility, which sits at the heart of information asymmetry. The next Part presents such a strategy.

### III. TRUTH BOUNTIES

Here we sketch a solution to misinformation: truth bounties. We develop a system built on voluntary pledges of conditional payments by speakers. The bounty would act as a promise or a bond that the speaker's statement is true. If the statement proves to be materially false, the speaker would lose the bounty. In offering a truth bounty, speakers would signal that they have confidence in the truthfulness of the information they share—so much confidence that they are willing to put money on the line.

Our solution is private and voluntary and operates based on contracts. Thus, it functions outside of public law, avoiding obstacles like the First Amendment. Unlike other reform proposals, ours works *ex ante*; speakers post bounties before their information disseminates. Our solution also operates at scale. Millions of communications could operate in the system we develop. More importantly, our solution addresses the central challenge facing information consumers in the digital age: credibility. Knowing that speakers have something to lose should make consumers more confident in their speech. By seeing how much speakers post, listeners can also learn something very important about the reliability of the information. Bounties send a finely tuned signal.

This Part explains our solution in detail, beginning with its conceptual roots.

#### A. Skin in the Game

Law often improves behavior through sanctions. People drive safely, respect others' property, and refrain from littering because doing otherwise will lead to some kind of negative consequence, whether imprisonment, a fine, or the payment of damages. Law forces people to put "skin in the game." They have something to lose—money, freedom—if they act wrongly, and this encourages them to act rightly. In the context of misinformation, many regulatory measures and proposals adopt this logic. They aim to sanction, in one way or another, the purveyors of lies.

Sometimes sanctions can do more than harm wrongdoers. They can help "rightdoers." A primary benefit of sanctions for misinformation is that they benefit honest purveyors of information by making their communications credible. To illustrate, suppose a seller advertises a product as having high quality and promises to deliver it tomorrow in exchange for a payment from the buyer today. The buyer would like the product if it is indeed high quality, but can she trust the seller? Talk is cheap, and the seller might send a low-quality product. Contract law overcomes the problem by threatening the seller with a sanction.<sup>129</sup> If she fails to deliver the product as she warranted, the buyer can sue for

---

<sup>129</sup> Criminal law has this feature too, although it is only used for severe transgression. *See Alvarez*, 567 U.S. 709 (2012).

damages. Contract law forces the seller to put skin in the game. If she breaks her promise, she will have to pay.<sup>130</sup> The threat of this sanction does not necessarily harm the seller. She *wants* buyers to trust her, and having skin in the game helps. The threat of liability makes her promises credible.<sup>131</sup>

The same idea operates in other areas of law. Manufacturers sell toasters, lawn mowers, medicines, and electric cars. They make representations about these products, such as “it has a range of 300 miles.” They want consumers to trust these statements and buy the products. Contract law regards these statements as warranties, compelling the manufacturer to pay if the product fails. In addition, consumer protection laws, including prohibitions on fraud and false advertising laws, make the representations trustworthy.<sup>132</sup> If the carmaker lies about the range, regulators will issue a fine, consumers and competitors will sue, or both. Law forces the manufacturers to put skin in the game, and this tends to help honest manufacturers. Having something to lose signals to consumers that they tell the truth.<sup>133</sup>

To summarize, law forces some people who supply information to put skin in the game, and this has two effects. It discourages some dishonest communications, and it makes the remaining communications credible.<sup>134</sup> Knowing that lies get punished increases trust in information.

We are interested in information in general, including information about politics and current events, not simply promises from sellers or representations about products. Can existing law force people who produce information in general to put skin in the game? Only to a limited degree. We have explained that defamation law only applies to information about specific people (“he robbed a bank”), not information in general (“Vaccines are fake”).<sup>135</sup> Furthermore, public figures like celebrities and politicians must prove actual malice to succeed in a defamation suit.<sup>136</sup> That high bar is hard to meet, meaning purveyors of false information often escape, and know they can escape, liability.<sup>137</sup>

---

<sup>130</sup> Note that warranties transcend (some) requirements, meaning that warranties are applicable also to users of the product who are not purchasing directly from the manufacturer. See Richard E. Speidel, *Warranty Theory, Economic Loss, and the Privity Requirement: Once More into the Void*, 67 B.U. L. REV. 9, 11 (1987).

<sup>131</sup> See ROBERT COOTER & THOMAS ULEN, *LAW & ECONOMICS* (6th ed. 2011).

<sup>132</sup> Consumer protections include products liability, which is complicated and may have many effects. See generally Keith N. Hylton, *The Law and Economics of Products Liability*, 88 NOTRE DAME L. REV. 2457 (2013).

<sup>133</sup> See, e.g., MARK A. GEISTFELD, *PRODUCT LIABILITY LAW*, 355 (2nd ed. 1999) (“There are good reasons for expecting that the prospect of liability gives sellers an incentive to invest in safer products.”).

<sup>134</sup> So long as courts cannot determine the truth with full accuracy, liability rules also chill some honest speech. For the law to produce a credible signal, consumers must believe that, on average, published statements are likely to be true. Their propensity to believe also depends on the costs of mistakes. For a full analysis, see Arbel & Mungan, *supra* note 31.

<sup>135</sup> *Id.*

<sup>136</sup> *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964).

<sup>137</sup> One can put skin in the game without law. Many newspapers try to report truthfully not only

In sum, having skin in the game should make producers of information more honest and trustworthy. But getting skin in the game is difficult. Defamation law applies only to some information producers. The First Amendment and other obstacles discussed earlier prevent law from doing much more. Reputation is limited in crowded and dynamic information environments. This does not mean the skin-in-the-game theory fails. Having something to lose should make statements more credible. However, law prevents us—lawmakers, regulators, ordinary citizens—from forcing most information producers to put skin in the game.

## B. The Voluntary Pledge

When sticks fail, carrots can do the trick. We propose a system built on voluntary, conditional payments by speakers. We call the conditional payment a truth bounty. The bounty would act as a promise backed by a bond that the speaker's statement is true. If the statement proves to be materially false, the speaker would lose the bounty. In posting a bounty, people would signal that they have confidence in the truthfulness of the information that they share. In choosing the size of the bounty, people could express not just their confidence, but also the *degree* of their confidence.

The bounty could not be too small, lest it be ineffective. If a person pledged, say, \$1, no one would take it seriously. This would not amount to putting skin in the game. On the other hand, the bounty could not be too large. If the system required a bounty of, say, \$10 million, many people would not have the resources to use the system, and the people who did would worry—and thus be overly cautious in their expression—to avoid any mistakes. They might lose a fortune without good cause. We will say more about mistakes later. For now, the point is simply that the optimal bounty presents a Goldilocks problem—not too small, not too large.

Who would set the amount of the bounty? In principle, the decision could be left to the speaker or set by the bonding system accepting the bounty.<sup>138</sup> The former would allow speakers to set an amount that corresponds to their level of confidence in the story. The higher the confidence level, the larger the bounty. It would also allow for flexibility based on resources. A \$1,000 bounty would mean more coming from a local newspaper or an independent investigative journalist than from a large company like CNN. On the other hand, having the system set the bounty could promote consistency and standardization. For

---

because they fear liability for defamation but because they fear a loss of reputation. Reputations for truthfulness pay off in terms of subscriptions, ads, and readership, and spreading misleading information would squander a good reputation. But reputation is not a panacea. *See generally* Yonathan Arbel, *Reputation Failure and the Limits of Market Discipline*, 54 WAKE FOREST L. REV. 1239 (2019).

<sup>138</sup> The Magnusson-Moss Act attempts to standardize the form of certain warranties, but much like our proposal does not require that any warranties be extended. 15 U.S.C. § 2301 et seq.

information consumers, those virtues might make truth bounties easier to understand.

In the simplest implementation, the speaker would deposit the truth bounty in escrow managed by a third party. This solution might not work well at scale. Mass speakers like the New York Times probably would not want to tie up so much capital in escrows. As an alternative, the speaker could post collateral, or a third party like an insurance company could underwrite the speaker's publications. This would limit the capital requirements for the speaker and enable them to use the system at scale.<sup>139</sup>

Information producers would advertise their truth bounties. This could be accomplished in various ways. For written communications, a natural way would involve the use of an icon. Imagine a news organization publishing a story on its website. An icon could appear next to the headline indicating that the story has a bounty. The icon would be standardized. Over time, users of the system would learn its meaning, just as people have learned to recognize trademarks like McDonald's arches and security icons deployed online to suggest secure commercial transactions. The icon would tell information consumers, whoever and wherever they are, that the news organization feels so confident about the story that it put money on the line. The following figure illustrates a possible implementation.



We started this section by contrasting sticks and carrots. We cannot force information producers to put skin in the game, but we can encourage them by offering something of value. In exchange for a truth bounty, they get the icon. For the reasons explained below, the icon symbolizes credibility. Seeing the amounts of money newspapers and other publishers spend on advertising their quality and reliability, there are good reasons to believe that many information producers would voluntarily pay for credibility.

<sup>139</sup> One might worry about the moral hazard and adverse selection inherent to any insurance scheme, but remember that insurers have various ways to ameliorate these problems. See Steven Shavell, *On the Social Function and the Regulation of Liability Insurance*, 25 GENEVA PAPERS ON RISK AND INSURANCE 166 (2000).

### C. Challenges and Fees

How does a truth bounty promote credibility? If the bounty is simply a loan—a third party holds the money for a while and then returns it—then the bounty and icon are meaningless. Speakers must bear risk. If a communication with a bounty attached turns out to be false, the speaker must lose the bounty. This is the key to the system.

To introduce risk, we propose a decentralized system of challenges. Suppose a reader sees a story with the icon indicating a bounty. Further suppose that the reader concludes that the story is false. Under a defamation regime, there is little the reader can do—unless he or she happens to be discussed in the story, and even then the legal hurdles are substantial. Under our proposed system, the reader could initiate a challenge. This is a critical feature of the system. *Any* member of the public could initiate a challenge to *any* communication with a bounty. Unlike defamation law, the system would not limit claims to the targets of specific allegations—it does not require any allegations at all.

Challenges could proceed in different ways, but for communications on the internet, a straightforward way would involve clicking on the icon. Doing so could bring challengers to a website. Information on the story in question—title, date, publisher, author, etc.—would load automatically, and the challenger could pursue her complaint.

Whether out of malice or ignorance, people could clog the system with meritless challenges. To mitigate this problem, the system could charge a challenge fee. The fee would force the challenger (whether a natural person or an entity) to put skin in the game. Like court fees, paying the challenge fee signals that the challenger has confidence in the merits of her claim.

As with the bounty, the challenge fee presents a balancing act. A small fee would fail to screen out meritless challenges, but a large fee could block even meritorious challenges. A similar problem arises when setting a court's filing fees.<sup>140</sup> One approach would be to make the challenge fee a single-digit percent of the bounty subject to some minimum. In any event, experience would inform the optimal amount.

After paying the fee, the challenger would have an opportunity to present her initial challenge. Basically, she would explain why she believes the story to be false. This process could take many forms. One approach would allow the challenger to present her argument in writing on the website and upload supporting files (images, audio, video). Afterwards the speaker would have an opportunity to rebut the challenger's initial complaint, again with text and possibly supportive files. This simple approach would not involve motions, oral arguments, or other trappings of a trial.

---

<sup>140</sup> See generally COOTER & ULEN, *supra* note 131, at 420-422.

The goal of this initial step would be to create a simple plausibility review. The challenge fee would also serve this function, but it could not prevent a deep-pocketed challenger from clogging the system. The bounty system—the entity collecting bounties and challenge fees and creating the icon—would conduct the plausibility review. The goal would be to screen out meritless claims: they are incoherent, have no evidence, make fantastical accusations, etc.

Assuming the challenger clears this bar, the case would proceed to arbitration.

#### D. Arbitration

The challenge would be resolved by arbitrators. The arbitral process could be informal and flexible, or it could resemble a trial. One could imagine other approaches between these poles. One could imagine presenting the parties with a menu of approaches from which they choose. We will not canvass specific possibilities but rather focus on some general features.

The arbitrators would be private actors. To avoid legal obstacles, nothing in the system requires state action. People would not get dragged into court against their will. No state or federal judge would assess the truthfulness of, say, someone's political speech, a possibility that raises serious concerns under the First Amendment. The system would be both voluntary and private, although some of its infrastructure might be provided by law.<sup>141</sup>

To build confidence in the system, arbitrators should be high-profile people with reputations for trustworthiness. As the system gains traction, the pool of arbitrators could broaden to encompass subject matter experts. Panels may include non-lawyers. Although helpful, training in law might not always be necessary.

To further build trust, the following mechanism could be used. The parties to a challenge could select an arbitrator from a pool. Each party would select one arbitrator, and the selected arbitrators would select the third, tie-breaking arbitrator. This system, which is common in arbitration, diminishes the risk of bias and appearances of bias. Selecting the decisionmaker creates legitimacy and makes it harder to complain about the outcome.

What exactly would the arbitrators decide? They could not decide whether a challenged communication is actually true or false. Deciding on the actual truth or falsity of a proposition raises deep challenges at the core of epistemology.<sup>142</sup> To demonstrate, readers probably assume (as do we) that the

---

<sup>141</sup> For example, arbitration awards are subject to legal review under certain, restrictive, conditions. See *Hall Street Associates, L.L.C. v. Mattel, Inc.*, 552 U.S. 576 (2008).

<sup>142</sup> See generally ALVIN I. GOLDMAN, *EPISTEMOLOGY AND COGNITION* (1986).

earth revolves around the sun and the moon is made of rocks, not cheese. But do we really know? Have you studied astronomy or tasted moon dust?<sup>143</sup>

Like judges and jurors in the formal legal system, the arbitrators would adopt a practical approach to truth guided by legal standards. Recall our news organization publishing a story on its website. In posting a bounty, the organization would not promise that every word is true, with no subsequent disproof possible. Such a promise would demand too much. A single error in a name, date, or location could cost the organization its bounty. Furthermore, some elements of a story might involve opinions, not facts. Opinions can be silly, uninformed, or whatever else, but adjudicating their falsity is either impossible or fraught with error.

For these reasons, posting a bounty would not commit the speaker to absolute truth. Rather, the speaker would commit to a standard. The optimal standard could be determined through experience (and perhaps vary by subject matter or industry). As a starting point, we suggest the following: “This information, taken as a whole, is materially accurate and not misleading.” The challenger’s burden would be to show that the information is materially inaccurate, or the information *is* misleading.

The term “materially” does important work. In general, small errors such as misspelled names or botched dates would not be material to the content of a story or other communication. Indeed, even defamation law does not consider these types of mistakes.<sup>144</sup> Thus, speakers could post bounties without fear of losing over a typo or silly mistake.

The materiality requirement and the “taken as a whole” language in our standard would require that the accuracy of a communication be assessed in a time-bound manner. The question is not whether the communication is accurate forevermore. The question is whether it was accurate at the time the bounty was posted given the information reasonably available to the speaker. To illustrate, suppose a person claimed in the year 1500 that the sun revolved around the earth. This was false, but the speaker could not reasonably have been expected to know that (Galileo came a century later). Liability for truth must account for what could have been known.

Under our proposed standard, a challenger could win by showing that a communication is “misleading.” A story can be accurate but misleading at the same time, as when a person truthfully reports the findings of one scientific study

---

<sup>143</sup> See, e.g., Richard Posner, *The Jurisprudence of Skepticism*, 86 MICH. L. REV. 827, 836 (1988) (“Many scientific theories, including natural selection and the “big bang,” are not verifiable by experimentation or any other method of exact observation; many have been proved false after having been universally accepted. . . many. . . are temporary or ad hoc constructs to explain phenomena that might be explained in other ways?”).

<sup>144</sup> KEETON ET AL., *supra* note 124, at 842 (explaining that in defamation law “it is not necessary to prove the literal truth of the accusation in every detail, and that it is sufficient to show that the imputation is substantially true, or, as it is often put, to justify the ‘gist,’ the ‘sting,’ or the ‘substantial truth’ of the defamation.”).

but then generalizes from it, without bothering to mention other, equally credible scientific studies that reach different conclusions. Thus, the terms “accurate” and “not misleading” in our proposed standard do independent work. Of course, demands for context are endless. The standard does not require one to provide every piece of illuminating information, only information that is critical to the proper interpretation of the statements.

The arbitrators would decide whether the challenger met the burden as laid out in the standard. In doing so, they would have to rely on a burden of proof. We suggest preponderance of the evidence. The question becomes: has the challenger shown by a preponderance of the evidence that the information, taken as a whole, is materially inaccurate or misleading? This will strike some readers as a low bar. We believe, however, that users of the system would welcome it (the standard mirrors defamation between private individuals). People and organizations would post bounties to build credibility. Compared to clear and convincing evidence or beyond a reasonable doubt, the preponderance of the evidence standard would make it easier for a challenger, thus increasing the credibility effect. At the same time, speakers need not worry about an onslaught of try-your-luck challenges due to three internal checks. The challenge fee and plausibility review offer the first two gatekeeping functions; the fee-shifting rule elaborated below is the third. A preponderance standard coupled with these screens would seem to strike a sensible balance. “The story is likely to be accurate and not misleading,” a reader might reason, “because otherwise someone would bring a challenge and win.”

Later we will provide some specific examples of how arbitration might work in practice. For now, we will conclude with two general points. First, a functional system would require decisions about many details: the formality of the process, motions and evidence, whether there are oral arguments, appeals, and so on. Those decisions raise an important tradeoff. Adopting a simple, informal process should tend to lower costs but cause more errors.<sup>145</sup> With few steps and limited evidence, errors would be inevitable, as when a true story is deemed false, or vice versa. Conversely, adopting a sophisticated, formal process should tend to increase costs but cause fewer errors. Obviously fewer errors would be better, but costs are not irrelevant. No one would use the system if it became too costly in time, money, or effort.<sup>146</sup>

Here is the second point. Our proposed standard would necessarily require arbitrators to exercise judgment. Does the alleged omission make the story materially inaccurate? Does that phrasing make the story misleading? Like ordinary people, different arbitrators would reach different judgments on those questions in some close cases. Other standards would inevitably raise versions of the same problem. No matter how the standard is phrased or the burden of proof defined, arbitrators would sometimes disagree, different arbitral panels

---

<sup>145</sup> COOTER & ULEN, *supra* note 131.

<sup>146</sup> *Id.*

would sometimes reach different decisions in the same dispute, and observers would sometimes disagree with arbitral decisions.<sup>147</sup>

We do not believe this problem is fatal. The formal legal system suffers from this problem, yet it appears to function well much of the time, especially in private disputes like the ones we imagine. Arbitral panels could write and publish short opinions explaining their decisions. As with judicial opinions, this might temper some backlash.<sup>148</sup> Having a strong suite of arbitrators, steeped in journalistic norms, would go a long way—and having the funds to recruit them, makes this option viable.

Importantly, speakers could hedge their communications to avoid the uncertainty inherent in close cases. Speakers who want to avoid finding themselves with a marginal case could either conduct more investigations to support their communications or, in the alternative, hedge and qualify the language they use. They could add qualifying language to their communications, explicitly distinguish opinions from facts, and otherwise make editorial choices that turn “borderline” communications into “clearly accurate and not misleading” communications. This would, of course, be a virtue.

### E. Rewards and Signals

After the arbitrators reached a decision, two things would happen: money would change hands, and the outcome would become public. To begin, we discuss the money. If the challenge succeeds, the challenger gets the bounty. To make it concrete, if the Tuscaloosa Today pledged \$10,000 on a story, and if a challenger convinced the arbitral panel that the story is materially inaccurate, the challenger would get the \$10,000. On the other hand, if the challenge failed, the bounty would remain intact.

The possibility of making money would encourage challengers, which would be important to the system, especially in its formative period. One could think about this feature of the system as outsourcing the search for truth. Challenges mean people are consuming information, recognizing the “bounty” icon, hunting for inaccuracies or misleading statements, and attempting to root them out. For some, the ability to refute falsities would be enough to earnestly participate. For others, the money would provide the incentive. For some, having the money come from the speaker would sweeten the deal. A fierce critic of Fox News would not only like to receive \$10,000, he might take pleasure in knowing that Fox News had to foot the bill.<sup>149</sup>

---

<sup>147</sup> See, e.g., Anthony Niblett & Albert H. Yoon, *Judicial Disharmony: A Study of Dissent*, INT. REV. L. ECON. 42, 60–71 (2015).

<sup>148</sup> William C. Vickrey et al., *Opinions as the Voice of the Court: How State Supreme Courts can Communicate Effectively and Promote Procedural Fairness*, NAT'L CENTER FOR STATE COURTS (2012) (Arguing that “rulings communicate not only to lawyers but also to the public and media and explain how courts resolve disputes and determine constitutional rights.”).

<sup>149</sup> Cf. Andrew T. Hayashi, *The Law and Economics of Animus*, 89 U. CHI. L. REV. 581 (2022).

So far, we have abstracted from the costs of running the system. In reality, operating the system, marketing its services, and performing arbitration would cost a lot. Much like the legal system, these costs must be funded. We believe the source of funding should not be the bounty itself. Instead of taking a portion of the bounty, funds could come from three other sources: initial fees paid by speakers that are independent of the bounty; the challenge fees; and, much like any other arbitration process, the parties themselves.

Parties to arbitration often pay for the arbitral process. However, we would augment that usual system with a critical design choice: fee shifting. Under the so-called American Rule, each party bears its own litigation costs. Under the English rule, which we endorse for this system, the loser pays for the process. In our context, if the challenger succeeds, the speaker loses the bounty and pays for the arbitration. If the challenger fails, the challenger pays the arbitrators.

Scholars have concluded that the English rule discourages weak challenges and encourages strong ones.<sup>150</sup> The intuition is straightforward. By increasing the cost to challengers whose challenges fail, fee shifting would push them to bring only strong challenges, meaning challenges likely to succeed. We believe that, for our system, this is a desirable feature. Given the inherent uncertainty in determining truth, it is desirable to have fewer costly arbitrations with close cases, while making clear-cut, winning cases easier to bring. Although the system we envision could not adjudicate with perfect accuracy the truth or falsity of every possible communication, it could effectively refute stories that are clearly false. That would be immensely valuable, and the English rule would help achieve that goal.

In addition to money changing hands, arbitration would result in publicity. The outcome of arbitration—whatever it is—must be publicized. Publicity could come through different channels. The winning party would naturally want to publicize their winning—“we successfully refuted the claim that X was taking bribes.” But some challengers would have smaller platforms than others, and from experience we know that parties can misreport the outcomes of proceedings. Hence, the system would function better with a formal, centralized method of reporting outcomes.

In the context of digital communications, this could happen through the icon—the same icon that indicates a communication has a bounty. The icon could be adjusted to send different messages. To begin, the icon could be, say, light green, indicating the information has a bounty but has not been challenged. After a challenge has been filed but before it has been resolved, the icon could

---

(explaining how the size of a fine *and* the recipient of the money can affect the fine’s power to deter).

<sup>150</sup> David A. Root, *Attorney Fee-Shifting in America: Comparing, Contrasting, and Combining the "American Rule" and "English Rule"*, IND. INT’L & COMP. L. REV. See also Theodore Eisenberg & Geoffrey P. Miller., *The English vs. The American Rule on Attorneys Fees: An Empirical Study of Attorney Fee Clauses in Publicly-Held Companies’ Contracts*, NYU L. ECON. RSCH. PAPER 10-52 (2010).

turn, say, yellow, indicating that a challenge is pending. Perhaps readers could, simply by clicking the icon, see the pleadings and the status of the dispute. After a challenge fails, the icon could turn dark green, indicating that the information has been successfully defended. This would make the information especially credible. If a challenge succeeds, the icon could turn red. In both of the latter cases, clicking on the icon could reveal the arbitral panel's opinion.

The system would track the records of speakers and challengers, and clicking (or hovering) on the icon could reveal this information. To illustrate, suppose a person reads a story online. The reader does not recognize the name of the author or publication, but she sees that the story has a bounty. Hovering over the icon reveals that the author has bounties on 100 stories, seven have been challenged, and all seven challenges have failed. This makes the reader especially confident about the accuracy of the story. Conversely, suppose the reader sees that the story has a bounty, and the yellow icon indicates that it has been challenged. Hovering over the icon reveals that the challenger has challenged 50 different stories and succeeded in 35 cases. This rightly makes the reader more skeptical of the story.

## F. Equilibrium and Use Cases

The system we envision would have broad reach. For digital information, anyone in the world with internet access and a credit card could attach a truth bounty to his or her communication. Anyone in the world with internet access and a little money could challenge such a communication. Although many details would need to be worked out, arbitration could proceed virtually, with no need for physical records or travel. This would greatly reduce costs and frictions growing from competing courts, jurisdictional questions, and legal standards that vary by state and nation.

If the system succeeded, a virtuous equilibrium would result. People everywhere would, upon seeing the icon, have greater confidence in the veracity of the information. They would know that the source of the information had skin in the game. *Not* seeing the icon would send a similarly helpful signal. Information without the bounty would be suspect. Because people could make money by challenging false or misleading information, relatively few people would attach bounties to such information, meaning relatively few challenges and arbitrations would take place, and that would tamp down costs.

The analogy of warranties is important here. Samsung, a large and well regarded company, voluntarily offers warranties with the sale of every fridge. The voluntary offer of warranties is a common marketing norm. It is used by large organizations like Samsung and small ones like the local tailor. Deciding to offer a warranty involves a financial risk. These firms understand that quite well. But they also understand the marketing potential of credibility. Truth bounties make it possible to warrant the truth. Let us now examine a few use cases, in order to make matters more concrete.

## 1. Media Publishing

A natural use for truth bounties, and the one we emphasize throughout, is media publishing. Under the proposed system, the editor of a publication could choose which stories to support with truth bounties. Staking money would have several benefits for the editor. Most obviously, a truth bounty elicits trust. Reader trust is the currency of mainstream journalism—and many newspapers pride themselves on the quality of their reporting. They advertise their quality, citing their own reporting standards as a reason to have confidence in their work and, by extension, to read or watch their communications.<sup>151</sup>

Truth bounties could also aid in product differentiation. Product differentiation is helpful in a competitive landscape and of particular importance for entrants (in this context, new media) who seek to establish themselves. In a sense, truth bounties are a form of advertising, but unlike general advertising which suffers from a cheap talk problem, truth bounties send a loud and clear signal.

It may seem paradoxical for a media company to voluntarily commit to the payment of money for stories, especially given the tenuous financial status of many newspapers today and the large volume of stories they run. But on careful consideration, truth bounties are more realistic than they may appear. In a sense, newspapers already put money behind many of their stories. There is a risk that somebody might take offense and bring a defamation lawsuit. While stories on issues of public import are protected, this protection is limited, as illustrated by the trial of former vice-presidential candidate Sarah Palin against the New York Times for an editorial linking her to a mass shooting.<sup>152</sup> Even if the Times was likely to win (it did), the expense of the trial must have been significant. Had the paper lost, the scope of liability could have been large and, perhaps worse, highly unpredictable. Truth bounties hedge risks—they stipulate amounts that are known in advance and can be controlled by the paper itself. And just like the implied stake demanded by defamation law, no payouts have to be made if the stories are true.

Finally, consider profits. Many media companies want to earn money. If the New York Times posted truth bounties, it might convince some skeptics to trust its reporting. If only a fraction of the millions of Fox News watchers bought a subscription, the Times could come out ahead. Likewise, if Fox News wanted to draw viewers and readers away from CNN and MSNBC, it could

---

<sup>151</sup> Efrat Nechushtai & Lior Zalmanson, *'Stay Informed', 'Become an Insider' or 'Drive Change': Repackaging Newspaper Subscriptions in the Digital Age*, 22 JOURNALISM 2035 (2021) (finding that, among the 55 top-circulated daily newspapers, every subscription pitch included information quality).

<sup>152</sup> Jeremy W. Peters, *Sarah Palin v. New York Times Spotlights Push to Loosen Libel Law*, N.Y. TIMES (Feb. 15, 2022), <https://www.nytimes.com/2022/01/23/business/media/sarah-palin-libel-suit-nyt.html>.

attach truth bounties to its stories. The ability to signal the quality of one's product is valuable, whether that product is a toaster, electric car, or information.

## 2. Campaign Speech

In January 2015, two Republican presidential candidates got into a heated debate on national television.<sup>153</sup> Rick Perry argued that Mitt Romney had supported health care mandates and that he was trying to cover up his past support for these policies.<sup>154</sup> “Rick, I’ll tell you what” Romney replied, turning to face his opponent and extending his hand “Ten Thousand bucks? Ten thousand dollar bet?” Romney was willing to stake this amount to prove his point, but he was rebuffed by Perry: “I’m not in the betting business.”<sup>155</sup>

This short exchange is revealing. It demonstrates, first, how much factual statements matter for political figures. Whether or not Romney supported healthcare was important for his candidacy.<sup>156</sup> Second, Perry had no skin in the game. Because the odds of Romney suing him for defamation and winning were very low, and because no one else could sue, the accusation was effectively cheap talk. Perry could just as easily have asserted that Romney was secretly a Democrat or a citizen of Russia. A more subtle point concerns the social reaction to this offer. Romney was roundly mocked for his response because “casually offering a \$10,000 bet” was a violation of a social norm—it made Romney appear “rich, elite, and out of touch.”<sup>157</sup>

Truth bounties could offer a helpful tool in politics. If such bounties had been in use, Romney could have staked \$10,000 (or more) behind his claim that he never supported healthcare mandates. To be specific, he could have written a statement after the debate, posted it online, and attached a truth bounty. To mitigate the social norms problem, perhaps Romney’s campaign or a political action committee, rather than Romney himself, could have staked the bounty. Anyone—not just Rick Perry—could have challenged Romney’s statement. Romney’s opponents would have relished the chance to disprove his claim and collect the bounty. Romney would have relished the chance to defend his claim

---

<sup>153</sup> For an exposition of this point, see Hemel & Porat, *supra* note 22.

<sup>154</sup> Amy Gardner & Philip Rucker, *Rick Perry Stumbles Badly in Republican Presidential Debate*, WASH. POST (Nov. 10, 2011), [https://www.washingtonpost.com/politics/republican-presidential-candidates-focus-on-economy/2011/11/09/gIQA5Lsp6M\\_story.html](https://www.washingtonpost.com/politics/republican-presidential-candidates-focus-on-economy/2011/11/09/gIQA5Lsp6M_story.html).

<sup>155</sup> *Mitt Romney’s ‘Out of Touch’ \$10,000 Bet*, THE WEEK (Jan. 8, 2015) <https://theweek.com/articles/479518/mitt-romneys-touch-10000-bet>.

<sup>156</sup> See Tom Cohen, *Romney Camp Seeks to Clarify Its Health Care Message*, CNN POLITICS (July 5, 2012), <https://www.cnn.com/2012/07/05/politics/health-care-romney/index.html> (asserting that Romney’s stance on the healthcare mandate during the 2012 election was a key issue due to the Republican desire to overturn the Affordable Care Act).

<sup>157</sup> *Mitt Romney’s ‘Out of Touch’ \$10,000 Bet*, THE WEEK (Jan. 8, 2015) <https://theweek.com/articles/479518/mitt-romneys-touch-10000-bet>. See also *Romney’s \$10k Gamble*, POLITICO (Dec. 11, 2011) <https://www.politico.com/story/2011/12/romneys-bet-wins-him-opening-on-attacks-070246>; Chris Cillizza & Aaron Blake, *Mitt Romney’s \$10,000 Mistake*, WASH. POST (Dec. 12, 2011), [https://www.washingtonpost.com/blogs/the-fix/post/mitt-romneys-10000-mistake/2011/12/11/gIQA9aEQpO\\_blog.html](https://www.washingtonpost.com/blogs/the-fix/post/mitt-romneys-10000-mistake/2011/12/11/gIQA9aEQpO_blog.html).

in a serious setting with professional arbitrators and a factual record. Best of all, voters would receive a meaningful signal of truthfulness amid the political noise.

Consider also the following example. In December of 1995, the Republican National Committee was under pressure to show that it supported expansions to Medicare. The Committee ran an advertisement depicting Chairman Haley Barbour holding an oversized cashier's check payable to "YOUR NAME HERE."<sup>158</sup> The ad, which was followed by the coupon reproduced below, offered \$1 million to anyone who could disprove the Republicans' assertion that they passed a balanced budget in 1995 and increased Medicare spending by 50%.

**MILLION DOLLAR MEDICARE CHALLENGE**

**HERE'S WHY YOU HAVE NO CHANCE FOR THE MILLION DOLLARS.**

The Republican National Committee will present a cashier's check for \$1 million to the first American who can prove the following statement is false: "In November 1995, the U.S. House and Senate passed a balanced budget bill. It increases total federal spending on Medicare by more than 50% from 1995 to 2002, pursuant to Congressional Budget Office standards." Responses must be postmarked by December 20, 1995.

**I BELIEVE YOU, HALEY.**  
 Let me know what I can do to help you get out the truth that Republicans are increasing Medicare spending.

**I DON'T BELIEVE YOU, HALEY.**  
 Enclosed is my analysis of why you are wrong.

Name \_\_\_\_\_  
 Address \_\_\_\_\_  
 City, State, Zip \_\_\_\_\_ Phone \_\_\_\_\_

Mail to: Haley Barbour, Million Dollar Medicare Challenge, RNC, 210 First Street, SE, Washington, DC 20003

**"Heard the one about Republicans 'cutting' Medicare?"**

**The fact is Republicans are increasing Medicare spending by more than half. I'm Haley Barbour, and I'm so sure of that fact I'm willing to give you this check for a million dollars if you can prove me wrong."**

Eighty different individuals tried to claim the prize.<sup>159</sup> Representative Gene Taylor, a Democrat, argued that the budget was not "balanced" as claimed; Mr. Charles Resor of Wilson, Wyoming, focused on the second part, claiming that the use of "increases" was fallacious, and the correct language should have been "would have increased."<sup>160</sup> The RNC responded to all of these claims with a form letter denying the prize. In the resulting litigation, the RNC argued that the advertisement was "parody," and, in the alternative, that the statement was not disproven.<sup>161</sup> The trial court rejected the first argument, finding that it was a

<sup>158</sup> USA TODAY (Dec. 13, 1995).

<sup>159</sup> Republican Nat'l Comm. v. Taylor, 299 F.3d 887, 889 (D.C. Cir. 2002).

<sup>160</sup> *Id.*

<sup>161</sup> *Id.* at 889-90.

binding offer, but it accepted the second argument and granted summary judgment.

This story has an important lesson about the utility and practicality of truth bounties. The fact that politicians found the need to buy credibility by staking funds is quite telling of an unmet market demand for credibility. At the same time, it exposes the risk of multiple claimants, with as many as 80 different claimants conducting investigations and seeking the \$1 million. On the other hand, the case is useful in assuaging practical concerns. The RNC summarily rejected all claims with a form letter; very few lawsuits were actually filed, and the few lawsuits filed were dismissed on summary judgment. Only two cases were appealed, and they were quickly dismissed. Importantly, the first-come, first-served language presumably solved over-participation and excessive litigation. The effect of the first-come language emphasizes the importance of a default rule that only the first in line can claim the bounty (although she is free to trade some of her rewards with others in exchange for better evidence).

Truth bounties could have many applications in politics. Here are two other examples. Candidates could attach bounties to their qualifications and background (“I served in the war, I am a citizen of the United States, I did not plagiarize my college thesis”). Super PACs and other groups could attach bounties to their political ads. This might be especially useful to “dark money” groups, which want people to believe their communications but do not want to reveal their donors. These uses and many others would not only help speakers, they would help listeners. Voters would find it easier to sort truth from lies.

### 3. Advertising

How do you sell mattresses? As many failed businesses have learned, having a good product is only one part of the battle, sometimes the easier one. Effective advertising is key. The problem for a mattress manufacturer is that pretty much every other manufacturer already promises “the best sleeping experience,” regardless of the quality of their product. Standing out is difficult.

Truth bounties could be used for commercial speech. A mattress manufacturer could make statements backed by a truth bond. For example, the manufacturer could claim that its mattress is made in the United States from top-quality latex, has been lab tested, or is clinically proven to reduce back pain. In all cases, the credibility benefits of having a truth bounty should materialize.

A case of desperados and welshers provides a striking illustration. Rudy Turilli operated a museum dedicated to the notorious desperado Jesse James. A central attraction of his museum was his theory that James did not die in a shootout. Instead, he assumed a secret alias and lived in what became Turilli’s museum until he passed away of old age. Turilli went on air, advertised his theory, and then offered \$10,000 (\$70,000 in 2022 USD) to “anyone who could

prove me wrong.” Turilli’s whole career was built around this assertion.<sup>162</sup> Unfortunately for him, the widow of James proved him wrong. Turilli refused to pay, claiming no contract was ever made and, if made, that his allegations were never properly refuted. The court, however, disagreed with him and ordered the payment of the bounty.

Other cases show the difficulty of collecting such bounties under existing law. Consider *Kolodziej v. Mason*.<sup>163</sup> During a television appearance a defense attorney claimed that his client couldn’t have traveled from his last known location to the scene of the crime in the relevant timeframe and was therefore innocent. The lawyer added “I challenge anybody to show me—I’ll pay them a million dollars if they can do it.”<sup>164</sup> An entrepreneurial law student accepted the challenge, replicating the trip and showing that it was manageable in time. The lawyer refused to pay, and the parties went to court. The judge ruled that there was no contract because the lawyer’s statement was indefinite and hyperbolic, comparable, the judge explained, to stating “I’ll be a monkey’s uncle.”<sup>165</sup> The judge refused to enforce a promise, made by a lawyer on national television, that contained a price and induced verifiable reliance by the plaintiff.

To be sure, false advertising law, misrepresentation doctrine, and Unfair, Deceptive, or Abusive Acts or Practices legislation already create some liability for advertisers, discouraging them from making misrepresentations about their products and services. But this liability is limited, either due to the high legal standard, the time and effort required for litigation, or the narrow scope of people who have standing to sue. By contrast, truth bounties come with a lower legal standard, and anyone could claim them.

Truth bounties are similar to product warranties. Indeed, the fact that many manufacturers provide warranties suggests the value and realism of truth bounties. Truth bounties are a generalization of this tried-and-true contractual mechanism.

These ideas cast light on the paradoxical nature of the puffery defense. Under the puffery doctrine, companies cannot be sued for false advertising if their statements can be interpreted as “mere puffery”—i.e., exaggeration, hyperbole, and other speech that is judged implausible. This is often interpreted as a pro-business rule because it shields firms from liability. But the converse is also true. Strong puffery defenses make all speech less credible, making it necessary for firms to invest more in advertising to win market share. Many of these investments are socially wasteful. If truth bounties could replace some of them, that would be another benefit.

---

<sup>162</sup> James v. Turilli, 473 S.W.2d 757, 761 (Mo. Ct. App. 1972) (“[D]efendant had virtually made a career out of his contention Jesse W. James was not killed in 1882 but lived many years thereafter as J. Frank Dalton.”).

<sup>163</sup> 774 F.3d 736, 740 (11th Cir. 2014).

<sup>164</sup> *Id.* at 739.

<sup>165</sup> *Id.* at 744.

## G. Sustaining Truth Bounties

We have explained how truth bounties could combat misinformation by distinguishing truthful from dishonest speech. By helping people make the distinction, the bounty system would benefit not only information consumers but also the many honest information producers who want to separate themselves from liars, swindlers, and propagandists. Once implemented, a bounty system would offer many advantages over other reform proposals touted today. But could it be sustained?

In game theory, it is common to examine systems by first assuming they work and then asking whether they will stop working once in place (i.e., asking whether they are an “equilibrium”). This way of thinking about the world is useful in many ways, one of which is that it highlights that however much effort is put into implementing policies, they can be undone quickly if they are not self-sustaining.

With this frame of mind, we can appreciate the power of truth bounties in creating incentives for actors to maintain the system. Consider a world where all the major newspapers use truth bounties extensively, politicians apply a bounty to their arguments on the campaign trail, and commercials by large advertisers usually include a bounty. Now consider a CEO of a news outlet who considers a money-saving reform: no bounties on any of company’s stories. It is easy to see why that would be tempting in the short run: no funds will have to be tied up, no bounties will have to be paid. But in the longer run, such a strategy would be destructive. Without truth bonds, readers would treat the news outlet with skepticism. The CEO would see readership and revenues drop.

This is not a mere thought experiment. We mentioned earlier strong evidence that labeling some stories as false leads readers to adjust their perceptions of *unlabeled* stories, considering them more credible.<sup>166</sup> The reverse would happen here. Deviating from a truth-bounty norm would be a clear red flag to readers that the source lacks credibility. Importantly, surveys of news consumption show that readers and viewers greatly care about source credibility when choosing which content to consume. A recent survey showed that 53% of US respondents said they prefer to pay for news than use free alternatives because paid news has “better quality.”<sup>167</sup>

We have explained that, once operational, a truth bounty system could sustain itself. *Making* it operational—developing the system in the first instance—is a separate and important challenge. One might even think it undermines our ideas. If the system we propose has so many advantages, why doesn’t it already exist? Here are two hypotheses. The technology necessary for a global truth bounty system has not been available for long, and the widespread focus on misinformation and search for solutions is relatively recent.

---

<sup>166</sup> Pennycook et al., *supra* note 114, at 4944-4957.

<sup>167</sup> *Understanding Value in Media: Perspectives from Consumers and Industry*, WORLD ECONOMIC FORUM (Apr. 2, 2020), <https://www.weforum.org/reports/value-in-media>.

#### IV. BOUNTIES AND THE FREEDOM OF SPEECH

We have described in detail the infrastructure necessary to create a system of truth bounties. Here we examine truth bounties on a higher level of abstraction. First, we consider the advantages of using a voluntary, contract-based approach to misinformation rather than a mandatory, tort-based approach. Second, we consider the relationship between truth bounties and equity. We do not believe the system would benefit the rich or harm the poor. Third, we consider the place truth bounties might occupy in contemporary speech debates and the opportunities they open for thinking about other alternative reforms.

##### A. Contracts vs. Torts

In the eyes of many, defamation law and other information regulations have failed.<sup>168</sup> Still, the standard response has been to double down.<sup>169</sup> Many reform proposals promise benefits, but some also have obvious and immediate risks: chilling reporting on controversial issues, increasing the cost of information production and hence information consumption, and making the government a truth arbiter in contentious domains.

Here we consider bounties as an alternative to the *proposed expansions* of current law. We do not make the absolute argument that bounties are superior to regulation, only that on the current margins, bounties are a better solution than the blunt expansion of defamation and similar laws.

In a world of exploding information sources, individuals cannot hope to vet all information hurtled at them. Instead, they opt to rely on basic filters and heuristics, such as only consuming information produced by sources perceived as reliable: a single news station, sources that tend to agree with one's pre-existing views (and thus have shown themselves to be reliable arbiters of truth, as the individual understands it), and homophilic attributes.<sup>170</sup> The danger is that such proxies can lead to echo chambers and polarization. By comparison, truth bounties offer a salient and direct signal of reliability. Rather than making the broad choice of CNN versus Fox News, one could pick and choose from all sources. The question is not "which station?" but rather "is there a bounty?"

A truth bounty is a tax on bullshit.<sup>171</sup> If the system took root, news without a bounty would find less demand and become less effective. The incentive to

---

<sup>168</sup> See, e.g., Verstraete, Bambauer & Bambauer, *supra* note 52, at 3 ("[M]any proposed solutions [to the problem of fake news] are unable to strike at the root of the problem[.]").

<sup>169</sup> See *supra* notes Part I.

<sup>170</sup> *Who Shared It?: How Americans Decide What News to Trust on Social Media*, AMERICAN PRESS INSTITUTE (Mar. 20, 2017), <https://www.americanpressinstitute.org/publications/reports/survey-research/trust-social-media/>.

<sup>171</sup> This is a paraphrase of Alex Tabarok, who studied bets in public discourse. Alex Tabarok, *A*

produce or spread misinformation would fall. To remain relevant, purveyors of misinformation could post bounties for their stories, but they would lose their money. Thus, bounties would discourage misinformation in two complementary ways: it would cost more to produce misinformation, and it would reach a smaller audience.

We would not expect the same effect from stricter defamation law. Although defamation law also imposes a tax on falsities, we explained earlier that it produces a noisy signal.<sup>172</sup> Knowing whether defamation law deters the reporting of false information requires one to know the implied stakes: can or will the victim sue, how strong is the case, what damages would the publisher owe, and so on. Truth bounties are more reliable because they are easier to interpret and claim. The stakes (i.e., the size of the bounties) are known and advertised in advance; they do not depend on legal details particular to this or that jurisdiction. Any member of the public could challenge any communication with a bounty. If the victim of a story lacks evidence, someone else might have it. Bounties work even when defamation does not apply, as when a story spreads lies but does not tarnish any reputations.

Defamation law, being a scion of tort law, also has a particular structural problem. In a nutshell, the deterrent effect of tort liability divorces the public value of information from the private harm.<sup>173</sup> Whether a newspaper should run a story will be affected by the scope of expected liability—but the scope of liability will be uncertain and *biased to protect the wealthy*. Expected liability when reporting about a random teacher from Oklahoma is lower than the expected liability of reporting about a socialite like Ghislaine Maxwell, who was found guilty of child sex trafficking. This might encourage the paper to report on the teacher. But the *social* value of reporting will often run in the opposite direction.<sup>174</sup>

Beyond the benefits to the general public, bounties also have advantages for publishers: credibility, predictability, cost, and coverage. Naturally, these advantages do not extend to all publishers: a scandalous tabloid would probably fare worse under a system of truth bounties. But for speakers who care about truth—either for its own sake or as a way of engendering trust—bounties could be extremely helpful.

One might think that the bounty system would disadvantage victims of defamation.<sup>175</sup> There are two distinct concerns, the magnitude of compensation

---

*Bet is a Tax on Bullshit*, MARGINAL REVOLUTION (Nov. 2, 2012, 7:35 AM), <https://marginalrevolution.com/marginalrevolution/2012/11/a-bet-is-a-tax-on-bullshit.html>.

<sup>172</sup> See discussion in Part 2.

<sup>173</sup> Because of this mismatch, newspapers will generally have suboptimal incentives to investigate and publish stories. See Matthew Gentzkow & Jesse M. Shapiro, *Competition and Truth in the Market for News*, 22 J. ECON. PERSP. 133 (2008); David J. Acheson & Ansgar Wohlschlegel, *The Economics of Weaponized Defamation Lawsuits*, 47 SW. L. REV. 335, 336 (2018).

<sup>174</sup> Indeed, this is why the Supreme Court created the exception for reports on issues of public concern, but this should be seen as patchwork.

<sup>175</sup> The true victims of defamation are all the people who were misled by it, and in the view of one

and its recipient. Under defamation law, the victim can get compensated for the full scope of her harm. In a bounty system, either the victim or any member of the public may claim the bounty, which may be smaller (or larger) than the true harm. These might seem like serious concerns. But we are not advocating for abolishing the existing system of liability. We are only arguing that it would be advantageous to pursue truth bounties instead of expansions to the liability regime. Unless defamation law goes away, which we do not advocate or anticipate, truth bounties would not cause victims to lose any rights.

Truth bounties offer another benefit to victims. To vindicate their good names under defamation law, victims must sue.<sup>176</sup> To vindicate their good names with a truth bounty, it is enough that *someone* brings a challenge. With many potential challengers drawing on a larger pool of evidence, the probability of vindication for a victim necessarily increases. Of course, no one could bring a challenge without a bounty, but this is another virtue of the system. Without a bounty attached, stories that defame people should garner less attention and be taken less seriously. The harms to the victim would decrease.

Truth bounties have a final, more structural benefit. Given the First Amendment, legislators and regulators cannot impose whatever sanctions they desire on false speech. To the extent bounties are voluntary, they sidestep these constraints. A writer, publisher, advertiser, scientist, politician or whoever else could choose to post a bounty, and someone could challenge it, without violating any constitutional norm.

## B. Equity and Access

Truth bounties aim to democratize the search for truth by enabling individuals to participate equally in the marketplace of ideas. A critical concern is to ensure equitable access for all individuals. On this score, we believe that truth bounties can outperform some alternative proposals, while complementing others. In contrast to defamation law, which imposes unpredictable and possibly significant financial risks on speakers and litigants,<sup>177</sup> and in contrast to reforms that would vest big platforms with editorial responsibilities that could disproportionately impact smaller outlets, truth bounties offer a more equitable approach.

To begin, consider information consumers. The system would aid all such consumers by helping them sort truths from falsehoods at zero cost. Insofar as poorer people have less education and fewer alternative tools for filtering out

---

of us, the subject of defamation has no privileged claim to priority in this sense. See Arbel, *supra* note 96. Still, in the analysis here we follow the conventional treatment of the victim.

<sup>176</sup> Many commentators argue that vindication is an important goal of defamation law. See Randall P. Bezanson, *The Libel Suit in Retrospect: What Plaintiffs Want and What Plaintiffs Get*, 74 CAL. L. REV. 189, 792 (1986); Randall P. Bezanson, *Libel Law and the Realities of Litigation: Setting the Record Straight*, 71 IOWA L. REV. 226, 228 (1985).

<sup>177</sup> See *supra* notes 125-27 and accompanying text.

misinformation, truth bounties would help them the most. We do not perceive any equity issues on the consumer side.

Now consider information producers. Using the truth bounty system would require some resources, which of course not everyone has. But it would require fewer resources than one might think. As discussed, communicators might be able to choose the size of the bounty. A \$100 bounty from a poor freelancer might mean more than a \$1,000 bounty from the Wall Street Journal. People can tailor the bounties to their resources. Furthermore, communicators only lose their bounty if their story turns out to be false or misleading. If their story is accurate, they keep it. Bounties would not operate forever; any particular bounty might be good for, say, one year. If no challenge succeeds during that time, the communicator gets the bounty back, possibly with interest. In short, honest communicators would not need lots of money to *spend*. They would need some money to *lend*.

To prevent frivolity, people challenging a bountied story would have to pay a fee. Paying such a fee would be challenging for poorer individuals. This problem is important but not unique to this context. The same issue arises with victims of accidents or defamation who cannot afford to sue. Some solutions are available, such as crowdfunding or third-party litigation funding. The latter option seems especially relevant in this context. The poor person A could partner with the resourced person B, with B paying the challenge fee, A disproving the story, and the parties splitting the bounty. Since many of the bountied stories would have a public interest component, it is possible that poor challengers could attract support from wealthy benefactors, NGOs, or public groups.

Truth bounties have another advantage for smaller players. Today many people segregate themselves into information silos, consuming information only from sources that they trust such as the Washington Post or Fox News.<sup>178</sup> This not only results in echo chambers; it advantages established publishers over smaller ones, making entry difficult. Truth bounties would allow entrants without established reputations to distinguish themselves by warranting the quality of their reporting. Warranties are a clear, battle-tested method of signaling reliability and attracting new clientele.<sup>179</sup>

We conclude by considering whether truth bounties would advantage wealthy actors. The main concern, we think, is that truth bounties would benefit *dishonest* wealthy actors who pursue strategic goals at the expense of truth. Rich players could, simply by attaching a bounty to their communication, send a signal that their communication is credible, even if it is actually false. In this way, a truth bounty system could magnify the power of lies by making them more persuasive. This is certainly a concern, but here are three reasons to think it is

---

<sup>178</sup> Linley Sanders, 'Trust in Media 2022', YouGov (Apr, 5, 2022) <https://today.yougov.com/topics/politics/articles-reports/2022/04/05/trust-media-2022-where-americans-get-news-poll>.

<sup>179</sup> See supra note 21 and accompanying text.

not serious. First, rich actors who post only small bounties may not gain much credibility. If a billionaire placed a \$1,000 bounty on his communication, would anyone take it seriously? To signal credibility, he might need a much larger bounty. Of course, a larger bounty means he will lose more when someone successfully challenges his communication, discouraging him from pursuing this strategy in the first place. Second, rich, dishonest actors might be willing to lose thousands or even millions of dollars here and there. We doubt, however, that many would be willing to lose those sums over and over by supporting one false communication after another. We would not expect a flood of bountied, false stories. Third, and most importantly, truth bounties have dynamic reputational effects. The system would maintain and publicize a record of bounties and challenges. People encountering a bountied story on the internet could quickly learn that the author had, for example, bountied 100 other stories, 11 of which were challenged, with just one challenge succeeding. Think of what this kind of information would mean for a wealthy, dishonest actor pursuing the strategy described above. He might bounty a false story to buy credibility, but consumers would see that he had bountied many other stories, most of which were challenged and successfully disproven. The actor's miserable track record would expose him and neutralize the credibility gains from his bounty.

In sum, no system to address something as complicated as truth will be fool-proof or offer completely equitable access to justice. But we believe that truth bounties could do well on these scores, particularly when compared to the status quo.

### C. Hands-On, Hands-Off, and the Invisible Hand

Should society regulate false information? Perhaps the ablest, sharpest champion of what we might call information *laissez-faire* is John Stuart Mill, who argued that “We can never be sure that the opinion we are endeavoring to stifle is a false opinion; and if we were sure, stifling it would be an evil still.”<sup>180</sup> This idea culminated in the metaphor of the marketplace of ideas, which was established when Justice Holmes in *Abrams v. United States* said that “the best test of truth is the power of the thought to get itself accepted in the competition of the market.”<sup>181</sup> The marketplace metaphor calls for a hands-off approach to information.

Technology today presents new challenges to the hands-off approach.<sup>182</sup> The digital revolution made it cheaper for low-quality information producers to spread their messages and drown the truth and easier for people to either get sucked into or comfortably maintain echo chambers around them. It also

---

<sup>180</sup> JOHN STUART MILL, ON LIBERTY 31 (1859).

<sup>181</sup> *Abrams v. United States*, 250 U.S. 616, 630 (1919).

<sup>182</sup> MINOW, *supra* note 3. Hasen suggests that recent advances pose a “clear and present danger” to people’s ability to judge the truth. HASEN, *supra* note 4, at 24.

undermined the traditional revenue models of high-quality journalism. As Professor Hasen argues, the problem is that speech is *too* cheap.<sup>183</sup>

Hasen and others who have argued along these lines might be right.<sup>184</sup> However, this does not point inevitably to regulation as the solution. Neither does exasperation with the flaws of regulation mean that we need to give up on any attempt to improve the information environment. Truth bounties offer an intermediate position between government regulation and *laissez-faire*. Truth bounties are autonomy-preserving and voluntary. They would operate through privately-run institutions. Their significance, then, goes beyond their effectiveness. They expand our choice set of how to design institutions for a modern society.

As an institution, truth bounties are respectful of autonomy because, unlike the one-size-fits-all approach of tort liability, they could allow each speaker to tailor her own potential exposure to liability. Having a choice of the extent to which we want to “put ourselves out there,” is an important aspect of our autonomy, but tort and criminal liability usually neglect this question.

Another appealing institutional feature is the respect for the agency of audiences, trusting their ability to discount statements that are backed by nothing more than words. In the bounty system, audiences have standing.<sup>185</sup> We emphasized throughout that audiences are reliably imperfect, but we think it is important not to pathologize them. The Supreme Court itself takes a rational view of audiences, making sure not to belittle them, which Lyrissa Lidsky powerfully defended by noting that:<sup>186</sup>

[A] State that indulges an irrationality assumption, or even a bounded rationality assumption, fails to respect the autonomy of its citizens, an autonomy upon which a self-governing democracy depends.<sup>187</sup>

Truth bounties are voluntary with speakers choosing their level of engagement. In practice, however, they might not feel optional. In a world where truth bounties are common, speakers will feel strong pressure to offer them. Self-regulation does not necessarily mean light, half-hearted, or sham regulation. It can be very powerful. Consider the many sellers offering return policies far more generous than they are legally required to offer, even at substantial cost to themselves.

Finally, while truth bounties are run by private organizations, they are backed by state institutions and infrastructure. Arbitration awards are backed by the enforcement power of the state, and the rules that govern truth bounty

---

<sup>183</sup> HASEN, *supra* note 4, at 24.

<sup>184</sup> See generally HASEN, *supra* note 4.

<sup>185</sup> See, e.g., *Norman v. Borison*, 418 Md. 630, 17 A.3d 697 (Md. Ct. Spec. App. 2011) (“the defamation of a company does not create a cause of action for its shareholders or owners”).

<sup>186</sup> Lidsky, *supra* note 19, at 805. See also *McConnell v. FEC*, 540 U.S. 93, 258 (2003) (Scalia, J., concurring in part and dissenting in part).

<sup>187</sup> See Lidsky, *supra* note 188, at 805.

institutions (the arbitral panels, the company accepting the bounties) are products of legislatures and possibly other, official lawmaking bodies.

## CONCLUSION

Misinformation threatens society. Many observers before us have proposed reforms meant to address this threat. While many of these proposals are thoughtful and valuable, many share a common oversight: they fail to take seriously credibility effects, that is, how they might affect not just speakers but also listeners. If a primary problem is that people believe the wrong sources, then we must be attentive to what forms people's credibility judgments. This is central to overcoming the key market failures, spillovers and information asymmetries, at the heart of free speech.

Having skin in the game begets trust. Perhaps more importantly, having skin in the game begets better information. "BS vendors," in the lively language of Nassim Taleb, cannot survive over time if they have to pay out of pocket.<sup>188</sup> This is why the Carboloc Smoke Ball company went bankrupt.<sup>189</sup>

Professor Hasen laments the rise of "cheap" speech, arguing that the "cheap speech era has threatened American democracy."<sup>190</sup> If this is the problem, the solution must involve making speech expensive.<sup>191</sup> What makes our approach novel is not this insight. Indeed, it underlies the many proposals to impose fines, sanctions, and tort liability on false speech. The novelty of our approach lies in developing a private mechanism for achieving this goal. The mechanism is new and occupies a position between top-down regulation and laissez-faire.

Truth bounties offer an autonomy-preserving alternative that can deeply impact our democracy and institutions. By pledging one's statements, one can broadcast confidence broadly and effectively. By allowing every member of the public to file a claim, truth bounties democratize the search for truth. By originating in private incentives, truth bounties can cover ground made immune to regulation by the First Amendment.

Implementing truth bounties is a challenge, but we believe it is feasible and—critically—self-sustaining. Once established in one domain, the institution can expand to others, the right kind of virality.

---

<sup>188</sup> Cf. NASSIM NICHOLAS TALEB, *SKIN IN THE GAME: HIDDEN ASYMMETRIES IN DAILY LIFE* (2018).

<sup>189</sup> A. W. B. Simpson, *Quackery and Contract Law: The Case of Carboloc Smoke Ball*, 14 J. LEG. STUD. 345, 368-75 (1985); *Carlill v. Carboloc Smoke Ball Co.*, 1 Q.B. 256 (Dec. 7, 1892).

<sup>190</sup> HASEN, *supra* note 4, at 26.

<sup>191</sup> Interestingly, Hasen does not take this approach. He advocates regulations ranging from funding disclosure rules to bans on targeting of election speech. HASEN, *supra* note 4, at 77-132.