

**Who Moderates the Moderators?: A Law & Economics Approach to Holding Online Platforms Accountable Without Destroying the Internet**

*Working Draft, November 9, 2021*

**Geoffrey A. Manne**  
**Kristian Stout**  
**Ben Sperry**

ICLE Working Paper 2021-11-09

## **Who Moderates the Moderators?: A Law & Economics Approach to Holding Online Platforms Accountable Without Destroying the Internet**

*Working Draft, November 9, 2021*

**Geoffrey A. Manne, Kristian Stout, & Ben Sperry**

### **Executive Summary**

A quarter-century since its enactment as part of the Communications Decency Act of 1996, a growing number of lawmakers have been seeking reforms to Section 230. In the 116<sup>th</sup> Congress alone, 26 bills were introduced to modify the law’s scope or to repeal it altogether. Indeed, we have learned much in the last 25 years about where Section 230 has worked well and where it has not.

Although the current Section 230 reform debate popularly—and politically—revolves around when platforms should be forced to host certain content politically favored by one faction (i.e., conservative speech) or when they should be forced to remove certain content disfavored by another (i.e., alleged “misinformation” or hate speech), this paper does not discuss, nor even entertain, such reform proposals. Rather, such proposals are (and should be) legal non-starters under the First Amendment.

Indeed, such reforms are virtually certain to harm, not improve, social welfare: As frustrating as imperfect content moderation may be, state-directed speech codes are much worse. Moreover, the politicized focus on curbing legal and non-tortious speech undermines the promise of making any progress on legitimate issues: The real gains to social welfare will materialize from reforms that better align the incentives of online platforms with the social goal of deterring or mitigating illegal or tortious conduct.

Section 230 contains two major provisions: (1) that an online service provider will not be treated as the speaker or publisher of the content of a third party, and (2) that actions taken by an online service provider to moderate the content hosted by its services will not trigger liability. In essence, Section 230 has come to be seen as a broad immunity provision insulating online platforms from liability for virtually all harms caused by user-generated content hosted by their services, including when platforms might otherwise be deemed to be implicated because of the exercise of their editorial control over that content.

To the extent that the current legal regime permits social harms online that exceed concomitant benefits, it should be reformed to deter those harms if such reform can be accomplished at sufficiently low cost. The salient objection to Section 230 reform is not one of principle, but of practicality: are there effective reforms that would address the identified harms without destroying (or excessively damaging) the vibrant Internet ecosystem by imposing punishing, open-ended legal liability? We believe there are.

First and foremost, we believe that Section 230(c)(1)'s intermediary-liability protections for illegal or tortious conduct by third parties can and should be conditioned on taking reasonable steps to curb such conduct, subject to procedural constraints that will prevent a tide of unmeritorious litigation.

This basic principle is not without its strenuous and thoughtful detractors, of course. A common set of objections to Section 230 reform has grown out of legitimate concerns that the economic and speech gains that have accompanied the rise of the Internet over the last three decades would be undermined or reversed if Section 230's liability shield were weakened. Our paper thus establishes a proper framework for evaluating online intermediary liability and evaluates the implications of the common objections to Section 230 reform within that context. Indeed, it is important to take those criticisms seriously, as they highlight many of the pitfalls that could attend imprudent reforms. We examine these criticisms both to find ways to incorporate them into an effective reform agenda, and to highlight where the criticisms themselves are flawed.

Our approach is rooted in the well-established law & economics analysis of liability rules and civil procedure, which we use to introduce a framework for understanding the tradeoffs faced by online platforms under differing legal standards with differing degrees of liability for the behavior and speech of third-party users. This analysis is bolstered by a discussion of common law and statutory antecedents that allow us to understand how courts and legislatures have been able to develop appropriate liability regimes for the behavior of third parties in different, but analogous, contexts. Ultimately, and drawing on this analysis, we describe the contours of our recommended duty-of-care standard, along with a set of necessary procedural reforms that would help to ensure that we retain as much of the value of user-generated content as possible, while encouraging platforms to better police illicit and tortious content on their services.

### **The law & economics of online intermediary liability**

An important goal of civil tort law is to align individual incentives with social welfare such that costly behavior is deterred and individuals are encouraged to take optimal levels of precaution against risks of injury. Not uncommonly, the law even holds intermediaries—persons or businesses that have a special relationship with offenders or victims—accountable when they are the least-cost avoider of harms, even when those harms result from the actions of third parties.

Against this background, the near-complete immunity granted to online platforms by Section 230 for harms caused by platform users is a *departure* from normal rules governing intermediary

behavior. This immunity has certainly yielded benefits in the form of more user-generated online content and the ability of platforms to moderate without fear of liability. But it has also imposed costs to the extent that broad immunity fails to ensure that illegal and tortious conduct are optimally deterred online.

The crucial question for any proposed reform of Section 230 is whether it could pass a cost-benefit test—that is, whether it is likely to meaningfully reduce the incidence of unlawful or tortious online content while sufficiently addressing the objections to the modification of Section 230 immunity, such that its net benefits outweigh its net costs. In the context of both criminal and tort law generally, this balancing is sought through a mix of direct and collateral enforcement actions that, ideally, minimizes the total costs of misconduct and enforcement. Section 230, as it is currently construed, however, eschews entirely the possibility of collateral liability, foreclosing an important mechanism for properly adjusting the overall liability scheme.

But there is no sound reason to think this must be so. While many objections to Section 230 reform—that is, to the imposition of any amount of intermediary liability—are well-founded, they also frequently suffer from overstatement or unsupported suppositions about the magnitude of harm. At the same time, some of the expressed concerns are either simply misplaced or serve instead as arguments for broader civil-procedure reform (or decriminalization), rather than as defenses of the particularized immunity afforded by Section 230 itself.

Unfortunately, the usual course of discussion typically fails to acknowledge the tradeoffs that Section 230—and its reform—requires. These tradeoffs embody value judgments about the quantity and type of speech that should exist online, how individuals threatened by tortious and illegal conduct online should be protected, how injured parties should be made whole, and what role online platforms should have in helping to negotiate these tradeoffs. This paper’s overarching goal, even more important than any particular recommendation, is to make explicit what these tradeoffs entail.

Of central importance to the approach taken in this paper, our proposals presuppose a condition frequently elided by defenders of the Section 230 status quo, although we believe nearly all of them would agree with the assertion: that there is actual harm—violations of civil law and civil rights, violations of criminal law, and tortious conduct—that occurs on online platforms and that imposes real costs on individuals and society at-large. Our proposal proceeds on the assumption, in other words, that there are very real, concrete benefits that would result from demanding greater accountability from online intermediaries, even if that also leads to “collateral censorship” of some lawful speech.

It is necessary to understand that the baseline standard for speech and conduct—both online and offline—is not “anything goes,” but rather self-restraint enforced primarily by incentives for deterrence. Just as the law may deter some amount of speech, so too is speech deterred by fear of reprisal, threat of social sanction, and people’s baseline sense of morality. Some of this “lost” speech will be over-deterred, but one hopes that most deterred speech will be of the harmful or, at least,

low-value sort (or else, the underlying laws and norms should be changed). Moreover, not even the most valuable speech is of infinite value, such that any change in a legal regime that results in relatively less speech can be deemed *per se* negative.

A proper evaluation of the merits of an intermediary-liability regime must therefore consider whether user liability alone is insufficient to deter bad actors, either because it is too costly to pursue remedies against users directly, or because the actions of platforms serve to make it less likely that harmful speech or conduct is deterred. The latter concern, in other words, is that intermediaries may—intentionally or not—facilitate harmful speech *that would otherwise be deterred (self-censored) were it not for the operation of the platform*.

Arguably, the incentives offered by each of the forces for self-restraint are weakened in the context of online platforms. Certainly everyone is familiar with the significantly weaker operation of social norms in the more attenuated and/or pseudonymous environment of online social interaction. While this environment facilitates more legal speech and conduct than in the offline world, it also facilitates more illegal and tortious speech and conduct. Similarly, fear of reprisal (i.e., self-help) is often attenuated online, not least because online harms are often a function of the multiplier effect of online speech: it is frequently not the actions of the original malfeasant actor, but those of neutral actors amplifying that speech or conduct, that cause harm. In such an environment, the culpability of the original actor is surely mitigated and may be lost entirely. Likewise, in the normal course, victims of tortious or illegal conduct and law enforcers acting on their behalf are the primary line of defense against bad actors. But the relative anonymity/pseudonymity of online interactions may substantially weaken this defense.

Many argue, nonetheless, that holding online intermediaries responsible for failing to remove offensive content would lead to a flood of lawsuits that would ultimately overwhelm service providers, and sub-optimally diminish the value these firms provide to society—a so-called “death by ten thousand duck-bites.” Relatedly, firms that face potentially greater liability would be forced to internalize some increased—possibly exorbitant—degree of compliance costs even if litigation never materialized.

There is certainly some validity to these concerns. Given the sheer volume of content online and the complexity, imprecision, and uncertainty of moderation processes, even very effective content-moderation algorithms will fail to prevent all actionable conduct, which could result in many potential claims. At the same time, it can be difficult to weed out unlawful conduct without inadvertently over-limiting lawful activity.

But many of the unique features of online platforms also cut *against* the relaxation of legal standards online. Among other things—and in addition to the attenuated incentives for self-restraint mentioned above—where traditional (offline) media primarily host expressive content, online platforms facilitate a significant volume of behavior and commerce that isn’t purely expressive. Tortious and illegal content tends to be less susceptible to normal deterrence online than in other contexts, as individuals can hide behind varying degrees of anonymity. Even users who are neither

anonymous nor pseudonymous can sometimes prove challenging to reach with legal process. And, perhaps most importantly, online content is disseminated both faster and more broadly than offline media.

At the same time, an increase in liability risk for online platforms may lead not to insurmountable increases in litigation costs, but to other changes that may be less privately costly to a platform than litigation, and which may be *socially* desirable. Among these changes may be an increase in preemptive moderation; smaller, more specialized platforms and/or tighter screening of platform participants on the front end (both of which are likely to entail stronger reputational and normative constraints); the establishment of more effective user-reporting and harm-mitigation mechanisms; the development and adoption of specialized insurance offerings; or any number of other possible changes.

Thus the proper framework for evaluating potential reforms to Section 230 must include the following considerations: To what degree would shifting the legal rules governing platform liability increase litigation costs, increase moderation costs, constrain the provision of products and services, increase “collateral censorship,” and impede startup formation and competition, all *relative to the status quo*, not to some imaginary ideal state? Assessing the marginal changes in all these aspects entails, first, determining how they are affected by the current regime. It then requires identifying both the direction and magnitude of change that would result from reform. Next, it requires evaluating the corresponding *benefits* that legal change would bring in increasing accountability for tortious or criminal conduct online. And, finally, it necessitates hazarding a best guess of the *net* effect. Virtually never is this requisite analysis undertaken with any real degree of rigor. Our paper aims to correct that.

### **A proposal for reform**

What is called for is a properly scoped reform that applies the same political, legal, economic, and other social preferences offline as online, aimed at ensuring that we optimally deter illegal content without losing the benefits of widespread user-generated content. Properly considered, there is no novel conflict between promoting the flow of information and protecting against tortious or illegal conduct online. While the specific mechanisms employed to mediate between these two principles online and offline may differ—and, indeed, while technological differences can alter the distribution of costs and benefits in ways that must be accounted for—the fundamental principles that determine the dividing line between actionable and illegal or tortious content offline can and should be respected online, as well. Indeed, even Google has argued for exactly this sort of parity, recently calling on the Canadian government to “take care to ensure that their proposal does not risk creating different legal standards for online and offline environments.”

Keeping in mind the tradeoffs embedded in Section 230, we believe that, in order to more optimally mitigate truly harmful conduct on Internet platforms, intermediary-liability law should develop a “duty-of-care” standard that obliges service providers to reasonably protect their users and others from the foreseeable illegal or tortious acts of third parties. As a guiding principle, we should

not hold online platforms vicariously liable for the speech of third parties, both because of the sheer volume of user-generated content online and the generally attenuated relationship between online platforms and users, as well as because of the potentially large costs to overly chilling free expression online. But we should place at least the same burden to curb unlawful behavior on online platforms that we do on traditional media operating offline.

Nevertheless, we hasten to add that this alone would likely be deficient: adding an open-ended duty of care to the current legal system could generate a volume of litigation that few, if any, platform providers could survive. Instead, any new duty of care should be tempered by procedural reforms designed to ensure that only meritorious litigation survives beyond a pre-discovery motion to dismiss.

Procedurally, Section 230 immunity protects service providers not just from liability for harm caused by third-party content, but also from having to incur substantial litigation costs. Concern for judicial economy and operational efficiency are laudable, of course, but such concerns are properly addressed toward minimizing the costs of litigation in ways that do not undermine the deterrent and compensatory effects of meritorious causes of action. While litigation costs that exceed the minimum required to properly assign liability are deadweight losses to be avoided, the cost of liability itself—when properly found—ought to be borne by the party best positioned to prevent harm. Thus, a functional regime will attempt to accurately balance excessive litigation costs against legitimate and necessary liability costs.

In order to achieve this balance, we recommend that, while online platforms should be responsible for adopting reasonable practices to mitigate illegal or tortious conduct by their users, they should not face liability for *communication* torts (e.g., defamation) arising out of user-generated content unless they fail to remove content they knew or should have known was defamatory. Further, we propose that Section 230(c)(2)'s safe harbor should remain in force and that, unlike for traditional media operating offline, the act of reasonable content moderation by online platforms should not, by itself, create liability exposure.

In sum, we propose that Section 230 should be reformed to incorporate the following high-level elements, encompassing two major components: first, a proposal to alter the underlying intermediary-liability rules to establish a “duty of care” requiring adherence to certain standards of conduct with respect to user-generated content; and second, a set of procedural reforms that are meant to phase in the introduction of the duty of care and its refinement by courts and establish guardrails governing litigation of the duty.

### **Proposed basic liability rules**

Online intermediaries should operate under a duty of care to take appropriate measures to prevent or mitigate foreseeable harms caused by their users' conduct.

Section 230(c)(1) should not preclude intermediary liability when an online service provider fails to take reasonable care to prevent *non-speech-related* tortious or illegal conduct by its users

As an exception to the general reasonableness rule above, Section 230(c)(1) should preclude intermediary liability for *communication* torts arising out of user-generated content unless an online service provider fails to remove content it knew or should have known was defamatory.

Section 230(c)(2) should provide a safe harbor from liability when an online service provider does take reasonable steps to moderate unlawful conduct. In this way, an online service provider would not be held liable simply for having let harmful content slip through, despite its reasonable efforts.

The act of moderation should not give rise to a presumption of knowledge. Taking down content may indicate an online service provider knows it is unlawful, but it does not establish that the online service provider should necessarily be liable for a *failure* to remove it anywhere the same or similar content arises.

But Section 230 should contemplate “red-flag” knowledge, such that a failure to remove content will not be deemed reasonable if an online service provider knows or should have known that it is illegal or tortious. Because the Internet creates exceptional opportunities for the rapid spread of harmful content, a reasonableness obligation that applies only *ex ante* may be insufficient. Rather, it may be necessary to impose certain *ex post* requirements for harmful content that was reasonably permitted in the first instance, but that should nevertheless be removed given sufficient notice.

### Proposed procedural reforms

In order to effect the safe harbor for reasonable moderation practices that nevertheless result in harmful content, we propose the establishment of “certified” moderation standards under the aegis of a multi-stakeholder body convened by an overseeing government agency. Compliance with these standards would operate to foreclose litigation at an early stage against online service providers in most circumstances. If followed, a defendant could provide its certified moderation practices as a “certified answer” to any complaint alleging a cause of action arising out of user-generated content. Compliant practices will merit dismissal of the case, effecting a safe harbor for such practices.

In litigation, after a defendant answers a complaint with its certified moderation practices, the burden would shift to the plaintiff to adduce sufficient evidence to show that the certified standards were not actually adhered to. Such evidence should be more than mere *res ipsa loquitur*; it must be sufficient to demonstrate that the online service provider should have been aware of a harm or potential harm, that it had the opportunity to cure or prevent it, and that it failed to do so. Such a claim would need to meet a heightened pleading requirement, as for fraud, requiring particularity.



Finally, we believe any executive or legislative oversight of this process should be explicitly scheduled to sunset. Once the basic system of intermediary liability has had some time to mature, it should be left to courts to further manage and develop the relevant common law.

Our proposal does not demand perfection from online service providers in their content-moderation decisions—only that they make reasonable efforts. What is appropriate for YouTube, Facebook, or Twitter will not be the same as what’s appropriate for a startup social-media site, a web-infrastructure provider, or an e-commerce platform. A properly designed duty-of-care standard should be flexible and account for the scale of a platform, the nature and size of its user base, and the costs of compliance, among other considerations. Indeed, this sort of flexibility is a benefit of adopting a “reasonableness” standard, such as is found in common law negligence. Allowing courts to apply the flexible common law duty of reasonable care would also enable the jurisprudence to evolve with the changing nature of online intermediaries, the problems they pose, and the moderating technologies that become available.

## **Who Moderates the Moderators?: A Law & Economics Approach to Holding Online Platforms Accountable Without Destroying the Internet**

*Working Draft, November 9, 2021*

**Geoffrey A. Manne, Kristian Stout, & Ben Sperry\***

### **I. Introduction / Overview**

A quarter-century since its enactment as part of the Communications Decency Act of 1996, a growing number of lawmakers have been seeking reforms to Section 230.<sup>1</sup> In the 116<sup>th</sup> Congress

---

\* The authors are, respectively: President and Founder, International Center for Law & Economics (ICLE) and Distinguished Fellow, Northwestern University Center on Law, Business, and Economics; Director of Innovation Policy, ICLE; and Associate Director for Legal Research, ICLE. ICLE is a nonprofit, nonpartisan research center that promotes the use of law & economics to inform public policy debates. ICLE has received financial support from numerous foundations, companies, and individuals, including firms with interests both supportive of and in opposition to the ideas expressed in this work. The ideas expressed here are the authors' own and do not necessarily reflect the views of ICLE's advisors, affiliates, or supporters. We would especially like to thank Neil Fried for formative discussions and copious intellectual inspiration, Matt Perault for valuable insights on and challenges to our ideas, and both for indispensable comments on earlier drafts. We would also like to thank Neil Chilson, Jesse Blumenthal, and the participants at the 2021 George Mason University Law & Economics Center Program on Economics & Privacy Scholars Research Roundtable for their helpful comments and discussions.

<sup>1</sup> Although the current Section 230 reform debate popularly—and politically—revolves around when platforms should be forced to host certain content politically favored by one faction (i.e., conservative speech) or when they should be forced to remove certain content disfavored by another (i.e., alleged “misinformation” or hate speech), this paper does not discuss, nor even entertain, such reform proposals. Rather, such proposals are (and should be) legal non-starters under the First Amendment. See, e.g., Mailyn Fidler, *The New Editors: Refining First Amendment Protections for Internet Platforms*, 2 NOTRE DAME L. SCH. J. EMERGING TECH. 241, 243 (2021) (“The editorial privilege protects the exercise of selection over the speech of others—curating speech. When platforms exercise selection over speech, they are protected as editors.”).

Indeed, such reforms are virtually certain to harm, not improve, social welfare: As frustrating as imperfect content moderation may be, state-directed speech codes are much worse. Moreover, the politicized focus on curbing legal and non-tortious speech undermines the promise of making any progress on legitimate issues: The *real* gains to social welfare will materialize from reforms that better align the incentives of online platforms with the social goal of deterring or mitigating illegal or tortious conduct. For further discussion of the well-canvassed First Amendment problems with these political speech efforts, see, for example, Ben Sperry, *tldr - The First Amendment & Section 230: Protecting Free Speech on the Internet*, INT'L CTR. L. & ECON. (Aug. 2020), <https://laweconcenter.org/wp-content/uploads/2020/08/tldr-The-First-Amendment->

alone, 26 bills were introduced to modify the law's scope or to repeal it altogether.<sup>2</sup> Indeed, we have learned much in the last 25 years about where Section 230 has worked well and where it has not.

Section 230 contains two major provisions: (1) that an online service provider will not be treated as the speaker or publisher of the content of a third party, and (2) that online service providers will not be liable for actions taken to moderate third-party content hosted by their services.<sup>3</sup> In essence, Section 230 has come to be seen as a broad immunity provision insulating online platforms from liability for virtually all harms caused by user-generated content hosted by their services, including when platforms might otherwise be deemed to be implicated because of the exercise of their editorial control over that content.

To the extent that the current legal regime permits social harms online that exceed concomitant benefits, it should be reformed to deter those harms, provided it can be done so at sufficiently low cost. The salient objection to Section 230 reform is not one of principle, but of practicality: are there effective reforms that would address the identified harms without destroying (or excessively damaging) the vibrant Internet ecosystem by imposing punishing, open-ended legal liability?

The debate over Section 230 reform is often framed as a binary choice: to maintain the statute as it is or to repeal it entirely.<sup>4</sup> But those are not, in fact, the only options. Various reform proposals each offer pieces of a useful approach, but few propose a holistic path forward. To more optimally mitigate truly harmful conduct on Internet platforms, we believe, first, that Section 230 immunity should be conditioned on a duty-of-care standard that obliges service providers to reasonably protect their users and others from the foreseeable illegal or tortious acts of third parties. But this alone would be deficient: adding an open-ended duty of care to the current legal system could generate a volume of litigation that few, if any, platform providers could survive. Thus, second, we believe that any new duty of care would need to be tempered by procedural reforms designed to ensure that only meritorious litigation survives beyond a motion to dismiss.

The crucial question is whether any proposed reform could pass a cost-benefit test—that is, whether it is likely to meaningfully reduce the incidence of unlawful or tortious online content while

---

[and-Section-230.pdf](#); Ben Sperry, *Committee Prepares to Grill Tech CEOs, but It Is the First Amendment That Could Get Torched*, TRUTH ON THE MARKET (Mar. 24, 2021), <https://truthonthemarket.com/2021/03/24/committee-prepares-to-grill-tech-ceos-but-it-is-the-first-amendment-that-could-get-torched/>.

<sup>2</sup> See Valerie Brannon & Eric Holmes, Congressional Research Service, *Section 230: An Overview*, at Summary, CONGRESSIONAL RESEARCH SERVICE (Apr. 7, 2021), available at <https://crsreports.congress.gov/product/pdf/R/R46751>.

<sup>3</sup> 47 U.S.C. § 230(c)(1)-(2).

<sup>4</sup> See, e.g., Eric Goldman, *An Overview of the United States' Section 230 Internet Immunity*, in THE OXFORD HANDBOOK OF ONLINE INTERMEDIARY LIABILITY 164 (Giancarlo Frosio, ed. 2020) ("There is another alternative: we could restore the offline publishers' liability rule to all online services and hold online services liable for all third-party content they publish.").

sufficiently addressing the objections to the modification of Section 230 immunity,<sup>5</sup> such that its net benefits outweigh its net costs. “The general problem remains one of selecting the mix of direct and collateral enforcement measures that minimizes the total costs of misconduct and enforcement.”<sup>6</sup> There is no reason to think this is impossible. While many objections to Section 230 reform are well-founded, they also frequently suffer from overstatement or insufficiently supported suppositions about the magnitude of harm.<sup>7</sup> At the same time, some of the expressed concerns are either simply misplaced or serve instead as arguments for broader civil-procedure reform (or decriminalization<sup>8</sup>), rather than as defenses of the particularized immunity afforded by Section 230 itself....

In what follows, we offer our analysis of these objections, as well as some proposals to reform Section 230 that, we believe, appropriately address the stated concerns and suggest a viable path forward.<sup>9</sup> These proposals are a working draft of what we believe may be the best way forward, but, more importantly, they reflect how this paper’s framework for assessing online intermediary liability can offer new insights and potential solutions to seemingly intractable problems. Many may challenge how *well* our suggestions navigate the relevant tradeoffs, but the overarching point of this exercise is to demonstrate *how* we should be negotiating the tradeoffs embedded in Section 230 and its reform.

Of central importance to the approach taken in this paper, our proposals presuppose a condition frequently elided by defenders of the Section 230 status quo, although we believe nearly all of them would agree with the assertion: that there is actual harm—violations of civil law and civil rights, violations of criminal law, and tortious conduct<sup>10</sup>—that occurs on online platforms and that

---

<sup>5</sup> See *infra* Part II.

<sup>6</sup> Reinier Kraakman, *Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy*, 2 J.L. ECON. & ORG. 53, 61 (1986). The same analysis underlies the assessment of indirect liability in other regimes. Under copyright law, for example, intermediaries and other third parties may be contributorily or vicariously liable. “[E]very mechanism for rewarding authors inevitably introduces some form of inefficiency, and thus the only way to determine the proper scope for indirect liability is to weigh its costs and benefits against the costs and benefits associated with other plausible mechanisms for rewarding authors.” William M. Landes & Douglas Gary Lichtman, *Indirect Liability for Copyright Infringement: An Economic Perspective*, 16 HARV. J.L. & TECH. 395, 410 (2003).

<sup>7</sup> See *infra* Part III.

<sup>8</sup> To take one important example, the harms to sex workers that have occurred since the passage of FOSTA/SESTA are really a function of the illegality of sex work, and not due primarily to the lack of Section 230 immunity. Where sex work is not illegal, as in parts of Nevada, websites that cater specifically to that demand can and do exist, even after the passage of the law. By the same token, in areas where the practice is legal, sex workers generally have much safer working conditions. See Annamarie Forestiere, *To Protect Women, Legalize Prostitution*, HARV. CIVIL RIGHTS-CIVIL LIBERTIES L. REV. (Oct. 1, 2019).

<sup>9</sup> For an outline of our proposed reforms, see *infra* Part IV.

<sup>10</sup> Throughout this paper, when we speak of harm, we mean *legally cognizable* harms in the sense that they are violations of civil statutes, the common law, or criminal codes. We aren’t calling for laws that would create new causes of action for such harms, nor contemplating the enforcement of laws that violate constitutional guarantees, most especially the First Amendment right to editorial discretion. Sometimes, as below in the discussion about Principle #3, harm can also mean awful but lawful content that platforms have a First Amendment right (further protected by Section 230) to moderate.

imposes real costs on individuals and society at-large. Our proposal proceeds on the assumption, in other words, that there are very real, concrete benefits that would result from demanding greater accountability from online intermediaries, even if that also leads to “collateral censorship” of some lawful speech.<sup>11</sup>

We use the word “censorship” intentionally. The clearest (but not the only) tradeoff in requiring online intermediaries to police more content is the loss of speech that may accompany it.<sup>12</sup> We also use the term for another reason: As suggested below, most defenders of the Section 230 status quo who fail to meaningfully address the potential benefits of more stringent restrictions on unlawful third-party content believe the costs of infringing free speech to be so high that they cannot possibly be justified by corresponding benefits. They are, in other words, free-speech absolutists. This is not our position, though we are staunchly defensive of free-speech rights and count the prospect of lost opportunities for user-generated speech as a significant potential cost of any limitation on Section 230 immunity.<sup>13</sup> Depending how speech is weighted in the calculus, some may conclude that the benefits of our proposed approach are not worth the costs. That is a tenable position. What is not tenable, however, is to *disregard* the benefits of reduced immunity, or to implicitly value speech as infinitely valuable, such that no benefit could ever be great enough to compensate for any reduction in speech.

Of course, even free-speech absolutists sometimes acknowledge that reform efforts may entail such a tradeoff. Indeed, in 2019, a group of 52 academics and scholars and 28 civil-society groups, including some of the staunchest defenders of online speech, proposed a set of “Principles for Lawmakers” to guide potential Section 230 reform.<sup>14</sup> These principles implicitly recognized the

---

<sup>11</sup> See Felix T. Wu, *Collateral Censorship and the Limits of Intermediary Liability*, 87 NOTRE DAME L. REV. 293 (2011); see also *infra* note 101 and accompanying text.

<sup>12</sup> See *infra* Part II.A

<sup>13</sup> It is worth noting that not all speech receives full First Amendment protection. Certain “low-value” speech like child pornography, revenge pornography, harassment, threats, incitement, and intimidation, fraud, defamation, and the like all receive little First Amendment protection for good reason. Regulation of unprotected speech is still limited by the doctrine of overbreadth and could even be struck down due to chilling effects on protected speech. See Richard Parker & David L. Hudson, Jr., *Overbreadth*, THE FIRST AMENDMENT ENCYCLOPEDIA (last updated 2017), <https://mtsu.edu/first-amendment/article/1005/overbreadth>; Frank Askin, *Chilling Effect*, THE FIRST AMENDMENT ENCYCLOPEDIA (2009), <https://www.mtsu.edu/first-amendment/article/897/chilling-effect>. Insofar as Section 230 immunity protects intermediaries from liability for illegal third-party speech beyond what the First Amendment would do, it has both benefits due to overcoming these chilling effects, and costs due to harms which are under-accounted for in many cases. Defenses of the Section 230 on speech grounds often ignore the costs of speech recognized in First Amendment doctrine. See, e.g., Eric Goldman, *Why Section 230 Is Better than the First Amendment*, 95 NOTRE DAME L. REV. REFLECTION 33, 46 (2019) (“Section 230 substantively protects more speech than the First Amendment, and the First Amendment will not adequately backfill any reductions in Section 230’s protections.”).

<sup>14</sup> *Liability for User-Generated Content Online: Principles for Lawmakers* (Jul. 11, 2019), available at <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2992&context=historical>. It should be noted that one of the authors of this paper, Geoffrey Manne, was one of the drafters (and a signatory) of the principles, and the organization for which all of the authors of this paper work, ICLE, was a signatory.

tradeoff, inasmuch as they acknowledged a theoretical path for reform and offered a framework to assess any proposed reforms. The top-level principles are:

- Principle #1: Content creators bear primary responsibility for their speech and actions;
- Principle #2: Any new intermediary-liability law must not target constitutionally protected speech;
- Principle #3: The law shouldn't discourage Internet services from moderating content;
- Principle #4: Section 230 does not, and should not, require "neutrality";
- Principle #5: We need a uniform national legal standard;
- Principle #6: We must continue to promote innovation on the Internet; and
- Principle #7: Section 230 should apply equally across a broad spectrum of online services.<sup>15</sup>

The goal of these principles is to preserve, as much as possible, the social gains that the Internet has provided, while directing any reform efforts toward targeting valid and well-defined harms. Thus, practical reforms that introduce harm-mitigation measures should also adopt appropriate constraints to adequately protect speech, encourage moderation, promote innovation, and avoid unnecessary administrative burdens.

Any reform efforts must begin with Principle #1: Content creators bear primary responsibility for their speech and actions. Obviously, platforms can and should be held responsible for their content when they are acting as content creators, as is the case under Section 230 today.<sup>16</sup> But holding platform *users* responsible means acknowledging that *platforms* may sometimes shield users from responsibility. It also means acknowledging that, in holding users responsible, it may be necessary to address the relationship between users and platforms, and not only the relationship between users and victims:

Third-party enforcement of any sort serves as a possible answer when deterrence fails because "too many" wrongdoers remain unresponsive to the range of practicable legal penalties. Direct

---

<sup>15</sup> *Id.*

<sup>16</sup> Actually, this is not entirely true. The language of Section 230 is so broad as to vitiate even straightforward vicarious liability where a publisher is responsible for content authored by its own employees. See *Blumenthal v. Drudge*, 992 F. Supp. 44, 51-52 (D.D.C. 1998) ("Because it has the right to exercise editorial control over those with whom it contracts and whose words it disseminates, it would seem only fair to hold AOL to the liability standards applied to a publisher or, at least, like a bookstore owner or library, to the liability standards applied to a distributor. But Congress has made a different policy choice by providing immunity even where the interactive service provider has an active, even aggressive role in making available content prepared by others."). See also Keith N. Hylton, *Property Rules, Liability Rules, and Immunity: An Application to Cyberspace*, 87 B.U. L. REV. 1, 36 (2007) ("Given that America Online's relationship with Drudge was similar to that between The New York Times and one of its columnists, the outcome of *Blumenthal* is difficult to defend. It is a long-settled matter of law that publishers are vicariously liable if they knowingly, recklessly, or negligently publish defamatory statements of the writers whose work they publish, and that newspaper publishers in particular are vicariously liable for the defamatory statements of their writer-employees.").

deterrence is the normal strategy for enforcing legal norms.... But of course direct deterrence sometimes fails for reasons that follow from its fundamental assumptions. It may fail because wrongdoers lack the capacity or information to make self-interested compliance decisions.... Yet, a more important source of failure is often the sheer cost of raising expected penalties high enough to deter wrongdoers.... Of course, these constraints on direct deterrence do not necessarily imply a need for supplemental enforcement measures. Alternative measures are justified only if they, in turn, can lower the total costs of direct enforcement and residual misconduct.<sup>17</sup>

Key to this acknowledgement is the basic rule that people respond to incentives. Conduct harmful to others is rarely deterred without external forces to provide those incentives. Sometimes, these forces take the form of inchoate social norms; sometimes, they are implicit threats of reprisal; sometimes, they are threats of law enforcement or civil liability. But arguably, the incentives offered by each of these forces is weakened in the context of online platforms. Certainly, everyone is familiar with the significantly weaker operation of social norms in the more attenuated and/or pseudonymous environment of online social interaction.<sup>18</sup> While this environment facilitates more legal speech and conduct than in the offline world, it also facilitates more illegal and tortious speech and conduct. Similarly, fear of reprisal (i.e., self-help) is often attenuated online, not least because online harms are often a function of the multiplier effect of online speech: it is frequently not the actions of the original malfeasant actor, but those of neutral actors amplifying that speech or conduct, that cause harm. In such an environment, the culpability of the original actor is surely mitigated and may be lost entirely. Likewise, in the normal course, victims of tortious or illegal conduct and law enforcers acting on their behalf are the primary line of defense against bad actors. But the relative anonymity/pseudonymity of online interactions may substantially weaken this defense.<sup>19</sup>

The point is that the baseline standard for speech and conduct is not “anything goes,” but rather, self-restraint enforced primarily by incentives for deterrence.<sup>20</sup> Just as the law may deter some amount of speech, so too is speech deterred by fear of reprisal, threat of social sanction, and people’s baseline sense of morality. Some of this “lost” speech will be over-deterred, but one hopes that most deterred speech will be of the harmful or, at least, low-value sort (or else, the underlying laws should

---

<sup>17</sup> Kraakman, *Gatekeepers*, *supra* note 6, at 56-57.

<sup>18</sup> See, e.g., Emily van der Nagel & Jordan Frith, *Anonymity, Pseudonymity, and the Agency of Online Identity: Examining the Social Practices of R/gonewild*, 20 FIRST MONDAY 3 (2015), <https://firstmonday.org/ojs/index.php/fm/article/view/5615/4346> (canvassing arguments on the harms associated with anonymity and pseudonymity online).

<sup>19</sup> See, e.g., David Lukmire, Note, *Can the Courts Tame the Communications Decency Act?: The Reverberations of Zeran v. America Online*, 66 N.Y.U. ANN. SURV. AM. L. 371, 402 (2010) (“*Zeran*’s interpretation that section 230(c)(1) forecloses ‘distributor’ liability eliminated any chance of recovery for plaintiffs in many Internet defamation cases. This consequence stems from the difficulty of identifying the original source of defamatory content on the Internet. Although the *Zeran* opinion claimed that plaintiffs could obtain redress, the ability to communicate anonymously on the Internet makes this reassurance illusory without distributor liability for several reasons.”).

<sup>20</sup> See Wu, *supra* note 11 and *infra* note 101 and accompanying text.

be changed). Where the incentives for self-restraint are mitigated, there will be more speech and conduct, and relatively more of it will be harmful or illegal. Any effort to reform Section 230 therefore should be designed to rebalance incentives to reflect broader social expectations more accurately.<sup>21</sup> The objective should not be to hold platforms liable for user-generated content, but, rather, to enlist platforms to mitigate harms when user-directed incentives are insufficient and to ensure that platforms do not prevent holding users appropriately responsible.

Principle #3—encouraging platform moderation—is, of course, consistent with this notion. The animating principle behind Section 230 was always to protect platforms from legal liability for their own efforts to deter undesirable online content.<sup>22</sup> Consistent with Principle #3, any reform efforts should work to accentuate, not diminish, platforms’ incentives to remove or prevent harmful or illegal content. But it must be noted that, because no moderation system is perfect, platform moderation of any sort necessarily entails prohibiting some perfectly legal or harmless content. The relevant question attending Section 230 reforms that encourage platforms to engage in *more* moderation is not whether this will deter some legal/harmless content (it will), but whether the marginal increase in the amount of legal/harmless content deterred is warranted.<sup>23</sup> Even with reforms that encourage more moderation, the combination of relatively weak social deterrence mechanisms in online spaces and the idiosyncratic nature of platforms’ moderation preferences may, of course, still leave us with more harmful online content than is socially optimal. Where current moderation practices comport well with social preferences, reforms should be careful not to impose unnecessary additional burdens. We believe that our proposals achieve this balance.

Reforms should encourage platforms to take steps to protect both their own users and non-users who may be harmed by illegal or tortious content on their platforms, just as “the threat of liability puts pressure on the owners of bars and restaurants to watch for any copyright infringement that might take place within their establishments; and the common law principle of vicarious liability obligates employers to monitor, train, and otherwise exercise control over the behavior of their employees.”<sup>24</sup> Under such an approach, online platforms need not be held liable for illegal or tortious user-generated content, but may be held to a duty of care requiring them to mitigate the extent of such content. The goal would be to develop a duty of care that platforms must exercise to

---

<sup>21</sup> Of course, if society believes the extent of deterrence (and thus of harm) online is preferable to the background level, that is fine. It should, however, be an explicit determination. The process of considering reforms to Section 230 should entail a discussion of this tradeoff; it should not assume it doesn’t exist.

<sup>22</sup> See JEFF KOSSEFF, *THE TWENTY-SIX WORDS THAT CREATED THE INTERNET* 83 (2019) (“Taken together, [subparts] (c)(1) and (c)(2) [of Section 230] mean that companies will not be considered to be the speakers or publishers of third-party content, and they will not lose that protection only because they delete objectionable posts or otherwise exercise good-faith efforts to moderate user content.”).

<sup>23</sup> See *infra* Part III.A.

<sup>24</sup> Doug Lichtman & Eric Posner, *Holding Internet Service Providers Accountable*, 14 SUP. CT. ECON. REV. 221, 224-25 (2006).



receive immunity, not to develop a federal tort system of vicarious liability. Thus, Section 230(c)(1)<sup>25</sup> should (and does) prevent holding a platform liable as the publisher of another's speech but should not (and arguably was never meant to) immunize a platform for its own failings to reasonably prevent unlawful conduct when it is the least-cost avoider.<sup>26</sup>

To be sure, imposing a duty-of-care standard would likely reduce *some* of the social benefits that platforms provide (e.g., innovation, expression, commerce, etc.). But if properly constructed, this approach should limit the loss of such benefits, while also limiting the social costs of unlawful or tortious online conduct.<sup>27</sup>

This paper proceeds as follows. Part II discusses the common objections raised in response to Section 230 reform efforts. Indeed, it is important to take those criticisms seriously, as they highlight many of the pitfalls that could attend imprudent reforms. We examine these criticisms both to find ways to incorporate them into an effective reform agenda, and to highlight where the criticisms themselves are flawed. Part III undertakes a law & economics analysis of platform moderation, introducing a framework to understand the tradeoffs faced by online platforms under differing legal standards with differing degrees of liability for the behavior and speech of third-party users. Part III also draws on common law and statutory antecedents that allow us to understand how courts and legislatures have been able to develop appropriate liability regimes for the behavior of third parties in different, but analogous, contexts. Finally, Part IV develops our recommended duty-of-care standard, along with a set of necessary procedural reforms that would help to ensure that we retain as much of the value of user-generated content as possible, while encouraging platforms to better police illicit and tortious content on their services.

## II. Common Objections to Section 230 Reform & Responses

A common set of objections to Section 230 reform has grown out of legitimate concerns that the economic and speech gains that have accompanied the rise of the Internet over the last three decades would be undermined or reversed if Section 230's liability shield were weakened. But these concerns are exaggerated in some dimensions, and in others stem from insufficiently developed conceptions of the litigation process and the proper metes and bounds of liability. As we discuss below, the law should be reformed to find liability where appropriate and where the costs of litigation to appropriately assign liability do not undermine the social utility of Internet services.

---

<sup>25</sup> 47 U.S.C. § 230(c)(1) states that “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”

<sup>26</sup> In law & economics parlance, a “least cost avoider” is the party to a conflict who can reduce the probability of a costly interaction happening at least cost. See Harold Demsetz, *When Does the Rule of Liability Matter?*, 1 J. OF LEG. STUD. 13, 28 (1972). See generally Ronald Coase, *The Problem of Social Cost*, 3 J. L. & ECON. 1 (1960).

<sup>27</sup> It is also important to note, as we discuss below, that many of the asserted costs from limiting immunity under Section 230 are likely far less significant than typically assumed or asserted. See *infra* Part II.

Few of the common objections to Section 230 reform grapple with this sort of cost-benefit assessment. Much more common is a presumption that any lost speech is of virtually infinite value, or that any change to the liability regime will generate virtually unbearable costs.<sup>28</sup> In no other area of the law is this true, and it is almost certainly not true of liability for online services, either.

Nonetheless, even if defenders of the status quo tend to overstate their position, the core concerns they express *are* legitimate and must be incorporated into any well-considered reforms.

### A. The moderators' dilemma

The immunity conferred by Section 230 was designed to overcome the so-called “moderators’ dilemma” that resulted from the decisions in 1991’s *Cubby Inc. v. CompuServe Inc.*<sup>29</sup> and 1995’s *Stratton Oakmont Inc. v. Prodigy Service Co.*<sup>30</sup> “Under the reasoning of *Stratton Oakmont*, online service providers that voluntarily filter some messages become liable for all messages transmitted, whereas [under *Cubby*,] providers that bury their heads in the sand and ignore problematic posts altogether escape liability.”<sup>31</sup> Facing a huge volume of third-party content and the risk of liability associated with performing any moderation of that content, online service providers would likely choose not to moderate at all. Alternatively, online platforms might simply decide that the liability risk was too large and opt to dramatically over-moderate or not host user-generated content at all—what some scholars term “collateral censorship.”<sup>32</sup>

Section 230 was drafted in large part (or perhaps, as some claim, *entirely*<sup>33</sup>) to avoid the moderator’s dilemma by protecting online intermediaries from liability for third-party speech, even

---

<sup>28</sup> See, e.g., Goldman, *Why Section 230 Is Better*, *supra* note 13, at 34 (“Section 230 provides significant and irreplaceable substantive and procedural benefits beyond the First Amendment’s free speech protections. Because the First Amendment does not backfill these benefits, reducing Section 230’s immunity poses major risks to online free speech and the associated benefits to society.”). Nowhere in Goldman’s argument on Section 230’s superiority to the First Amendment does he attempt to defend *why* the protection of speech beyond that conferred by the First Amendment is beneficial. He merely assumes it and believes his readers will do the same. As a matter of polemics, this is powerful, but as a matter of scholarship, this is lacking.

<sup>29</sup> 776 F. Supp. 135 (S.D.N.Y. 1991).

<sup>30</sup> 1995 WL 323710 (N.Y. Sup. Ct. 1995).

<sup>31</sup> Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC, 521 F.3d 1157, 1163 (9<sup>th</sup> Cir. 2008). For a detailed and insightful discussion of the implications of these two cases, see KOSSEFF, TWENTY-SIX WORDS, *supra* note 22, at 48-72.

<sup>32</sup> See Wu, *supra* note 11, at 295-96. See also Jack M. Balkin, *Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2298 (1999).

<sup>33</sup> See *Roommates.com*, 521 F.3d at 1164 n.12 (“While the Conference Report refers to this as ‘[o]ne of the specific purposes’ of section 230, it seems to be the principal or perhaps the only purpose. The report doesn’t describe any other purposes, beyond supporting ‘the important federal policy of empowering parents to determine the content of communications their children receive through interactive computer services.’”) (citation omitted).

when they engage in moderation that may make them aware of the presence of problematic content.<sup>34</sup> As Judge Alex Kozinski remarked in the *Roommates.com* case:

In passing section 230, Congress sought to spare interactive computer services this grim choice by allowing them to perform some editing on user-generated content *without thereby becoming liable* for all defamatory or otherwise unlawful messages that they didn't edit or delete. In other words, Congress sought to immunize the *removal* of user-generated content, not the *creation* of content.<sup>35</sup>

This last point is important, though somewhat misstated by Judge Kozinski. Even without Section 230, intermediaries would almost never be directly liable for the *removal* of user-generated content.<sup>36</sup> Rather, the removal of some content might (as in *Stratton Oakmont*) lead to liability for the *non-removal* of *other* unlawful content, on the theory that engaging in any moderation means the intermediary is either presumed to have contributed to (i.e., exercised editorial discretion over) the posting of all content or else should reasonably be aware of the presence of other, related unlawful content.

It is noteworthy, however, that the immunity conferred under Section 230(c)(1) is evidently related to communication torts (most significantly, defamation) as opposed to either conduct-based

---

<sup>34</sup> S. REP. NO. 104-230, at 194 (1996) (“This section provides ‘Good Samaritan’ protections from civil liability for providers or users of an interactive computer service for actions to restrict or to enable restriction of access to objectionable online material. One of the specific purposes of this section is to overrule *Stratton-Oakmont v. Prodigy* and any other similar decisions which have treated such providers and users as publishers or speakers of content that is not their own because they have restricted access to objectionable material. The conferees believe that such decisions create serious obstacles to the important federal policy of empowering parents to determine the content of communications their children receive through interactive computer services.”).

<sup>35</sup> *Roommates.com*, 521 F.3d at 1163 (emphasis added). See also Testimony of Former U.S. Rep. Chris Cox (Author and Co-Sponsor with Sen. Ron Wyden, Section 230), Hearing before S. Subcomm. on Commc’ns., Tech., and the Internet on “The PACT Act and Section 230: The Impact of the Law that Helped Create the Internet and an Examination of Proposed Reforms for Today’s Online World” (116<sup>th</sup> Cong., Jul. 28, 2020) at 12, available at <https://www.commerce.senate.gov/services/files/BD6A508B-E95C-4659-8E6D-106CDE546D71> (“What Section 230 added to the general body of law was the principle that an individual or entity operating a website should not, in addition to its own legal responsibilities, be required to monitor all of the content created by third parties *and thereby become derivatively liable* for the illegal acts of others.”) (emphasis added) [hereinafter “Cox, PACT Act Testimony”].

<sup>36</sup> The primary exception, presumably, would be when removing content was a violation of a contract with the plaintiff precluding such removal. See *Damaa, LLC v. Google, Inc.*, No. 15-cv-03221-RMW, at \*14 (N.D. Cal. Nov. 2, 2016) (“Plaintiff’s claim for breach of the implied covenant of good faith and fair dealing, however, is not precluded by § 230(c)(1) because it seeks to hold defendants liable for breach of defendants’ good faith contractual obligation to plaintiff, rather than defendants’ publisher status.”). But this is not remotely the typical case, nor is it the sort of conduct that Section 230 was intended to immunize. See, e.g., *King v. Facebook, Inc.*, No. 19-cv-01987-WHO, at \*7 (N.D. Cal., Sep. 5, 2019) (“[In *Damaa*], the defendant took actions against the economic interest of plaintiff and in favor of defendants’ business partner. That did not happen here, and the rationale of *Damaa* does not save [Plaintiff’s] currently alleged claims.”). One could also imagine fairly convoluted scenarios in which removing certain content directly renders some piece of remaining content illegal (say, perhaps, by removing a mandatory disclosure). But in such cases the problem would ultimately lie with the content that was not removed, rather than the content that was removed.

causes of action (i.e., relating to physical harm to persons or property) or actions vindicating civil rights.

In recognition of the speed with which information may be disseminated and the near impossibility of regulating information content, Congress decided not to treat providers of interactive computer services like other information providers such as newspapers, magazines or television and radio stations, all of which may be held liable for publishing or distributing obscene or defamatory material written or prepared by others. While Congress could have made a different policy choice, it opted not to hold interactive computer services liable for their failure to edit, withhold or restrict access to offensive material disseminated through their medium.<sup>37</sup>

Courts have mostly read Section 230(c)(2)<sup>38</sup> out of the law by holding that Section 230(c)(1) protects the “exercise of a publisher’s traditional editorial functions—such as deciding whether to publish, withdraw, postpone or alter content”<sup>39</sup>—which, after 1997’s *Zeran v. America Online* decision, has also been interpreted to include not treating online intermediaries as distributors.<sup>40</sup>

*Zeran* also pulled into Section 230’s ambit a wide swath of tort law not obviously contemplated by the statute’s defamation-based language or even its intent.<sup>41</sup> Instead of limiting its holding to the sort of communication torts (such as defamation) raised by the facts of the case—and, as noted, arguably contemplated by the statute’s use of defamation-related terms of art—*Zeran* used

---

<sup>37</sup> *Blumenthal v. Drudge*, 992 F. Supp. 44, 49 (D.D.C. 1998). See 47 U.S.C. § 230(c)(1) (“No provider or user of an interactive computer service shall be treated as the ***publisher or speaker*** of any information provided by another information content provider.”) (emphasis added). See also Goldman, *Why Section 230 Is Better*, *supra* note 13, at 36 (“Defamation is Section 230’s paradigmatic application.”); Danielle Keats Citron & Mary Anne Franks, *The Internet as Speech Machine and Other Myths Confounding Section 230 Reform*, 3 U. CHI. LEG. FORUM 45, 57 (2020) (“The text of Section 230 reinforces [the centrality of speech] through the use of the terms ‘publish,’ ‘publishers,’ ‘speech,’ and ‘speakers’ in 230(c), as well as the finding that the ‘Internet and other interactive computer services offer a forum for a true diversity of political discourse, unique opportunities for cultural development, and myriad avenues for intellectual activity.’”).

<sup>38</sup> Under 230(c)(2) “[n]o provider or user of an interactive computer service shall be held liable on account of—(A) any action voluntarily taken in good faith to restrict access to or availability of material” for a variety of reasons, including material that is merely “objectionable.” See 47 U.S.C. § 230(c)(2).

<sup>39</sup> *Zeran v. Am. Online, Inc.*, 129 F.3d 327, 330 (4th Cir. 1997).

<sup>40</sup> See *id.* at 331-34. Under traditional defamation law principles “primary” publishers of libelous content—those that exercise editorial control over the content they publish—are treated differently than “secondary” publishers that merely distribute already published content (i.e., distributors). The *Zeran* court reasoned that because distributors are a “species” of publisher under defamation law, and because traditional distributor liability would impair what it identified as the statute’s purpose, the word “publisher” in 230(c)(1) would be read to include both. See *id.* at 333 (“Because the probable effects of distributor liability on the vigor of Internet speech and on service provider self-regulation are directly contrary to Section(s) 230’s statutory purposes, we will not assume that Congress intended to leave [distributor liability] intact.”).

<sup>41</sup> See Lukmire, *supra* note 19, at 395 (“*Zeran*’s most unsettling move was broadening the reach of section 230(c)(1) beyond the boundaries of defamation law to cover a broad range of claims. Instead of interpreting section 230 as precluding only claims against primary publishers of third-party content (which *Zeran* advocated) or, even more broadly, as a bar against all defamation or defamation-type lawsuits related to disseminating third-party content (which the court’s collapse of distributor liability into the ‘publisher’ category might suggest), the court employed the more general terms ‘tort-based lawsuits’ and ‘tort liability’ to describe the scope of the safe harbor.”).

expansive language to describe the scope of Section 230's immunity: "By its plain language, § 230 creates a federal immunity to *any cause of action* that would make service providers liable for information originating with a third-party user of that service."<sup>42</sup> Although originally framed as relevant primarily to holding online message boards liable as republishers of defamatory material,<sup>43</sup> Section 230 immunity has subsequently grown far beyond the bounds of defamation law and "has led to a far broader immunity shield than would be implied by common law tort doctrine."<sup>44</sup> Section 230 has been invoked successfully in cases with causes of action that include negligence, deceptive trade practices, unfair competition, false advertising, common-law privacy torts, tortious interference with contract or business relations, and intentional infliction of emotional distress, among others.<sup>45</sup> Indeed, some have argued that "the broad construction of the CDA's immunity provision adopted by the courts has produced an immunity from liability that is far more sweeping than anything the law's words, context, and history support."<sup>46</sup>

---

<sup>42</sup> *Zeran*, 129 F.3d at 330 (emphasis added).

<sup>43</sup> See Brannon & Holmes, *supra* note 2, at 6-7.

<sup>44</sup> Keith N. Hylton, *Property Rules, Liability Rules, and Immunity: An Application to Cyberspace*, 87 B.U. L. Rev. 1, 37 (2007). See also Ryan J.P. Dyer, *The Communication Decency Act Gone Wild: A Case for Renewing the Presumption Against Preemption*, 37 SEATTLE U.L. REV. 837, 862 (2014) ("Section 230 of the CDA was enacted to remove the disincentive for online intermediaries to take good faith efforts to monitor and remove offensive content from their websites. Specifically, Congress meant to remove traditional forms of publisher liability and the accompanying legal exposure in the context of defamatory and pornographic content posted by third parties.... Unfortunately, early courts interpreting section 230 over-read the scope of immunity provided by the provision and erroneously broadened the range of civil and criminal liability schemes subject to preemption. The negative consequences of this misreading are increasingly felt as more and more criminal activity migrates to the Internet, and the online intermediaries that knowingly host such activity are held immune from traditional modes of checking such lawlessness.").

<sup>45</sup> See Goldman, *Why Section 230 Is Better*, *supra* note 13, at 36-37.

<sup>46</sup> Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 FORDHAM L. REV. 401, 408 (2017). It should be noted, however, that in recent testimony before the Senate, former U.S. Representative Chris Cox—one of the principal drafters of Section 230 along with (then-Representative, now-Senator) Ron Wyden—stated that:

[I]n enacting Section 230, it was not our intent to create immunity for criminal and tortious activity on the internet. To the contrary, our purpose (and that of every legislator who voted for the bill) was to ensure that innocent third parties will not be made liable for unlawful acts committed wholly by others.

Cox, *PACT Act Testimony*, *supra* note 35, at 8. In that same hearing, Cox also spoke largely approvingly of the development of Section 230 jurisprudence over the last twenty-five years. See *Id.* at 10-11. While it is certainly possible that Cox is accurately reporting his purpose "and that of every other legislator who voted for the bill," it is impossible to discern in the legislative discussions surrounding the bill an intention to construe the law so broadly. See 141 Cong. Rec. H8468-72 (Aug. 4, 1995). A key factor in this ambiguity is surely the use of the term "liability for content created by their users." Cox, *PACT Act Testimony*, *supra* note 35, at 3. As we discuss throughout this paper, there is a difference between indirect or vicarious liability under the causes of action that make certain content unlawful, and liability under a distinct cause of action stemming from an intermediary's violation of its duty of care in its treatment of such content. As Cox suggests in his testimony, the concern with indirect liability in the statute was focused on the specter of liability created by *Stratton-Oakmont* for underlying causes of action applied to intermediaries simply for engaging in the act of content moderation: "What Section 230 added to the general body of law was the principle that an individual or entity operating a website should not, in addition to its own legal

Of course, “the law’s words, context, and history” are not dispositive. As Judge Easterbrook wrote for the 7<sup>th</sup> Circuit in a case interpreting the scope of content reached by Section 230:

Section 230(c)(1) is general. Although the impetus for the enactment of § 230(c) as a whole was a court’s opinion holding an information content provider liable, as a publisher, because it had exercised some selectivity with respect to the sexually oriented material it would host for customers, *a law’s scope often differs from its genesis*. Once the legislative process gets rolling, interest groups seek (and often obtain) other provisions.

Congress could have written something like: “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any *sexually oriented material* [emphasis in original] provided by another information content provider.” That is not, however, what it enacted. Where the phrase “sexually oriented material” appears in our rephrasing, the actual statute has the word “information.” *That covers ads for housing, auctions of paintings that may have been stolen by Nazis, biting comments about steroids in baseball, efforts to verify the truth of politicians’ promises, and everything else that third parties may post on a web site; “information” is the stock in trade of online service providers.*<sup>47</sup>

Even then, however (and despite the hyperbolic “and everything else that third parties may post on a web site”), it’s a stretch to interpret Section 230 to preclude responsibility for the transmission of all tortious or criminal conduct.

Indeed, as the court writes in *Chicago Lawyers’ Committee*: “Subsection (c)(1) does not mention ‘immunity’ or any synonym. Our opinion in *Doe* explains why § 230(c) as a whole cannot be understood as a general prohibition of civil liability for website operators and other online content hosts.”<sup>48</sup> Such a broad immunity would defeat the underlying intention of the provision. As Judge Easterbrook wrote in his earlier *Doe* opinion (referenced in the above quote from *Chicago Lawyers’ Committee*):

If this reading is sound, then § 230(c) as a whole makes ISPs indifferent to the content of information they host or transmit: whether they do (subsection (c)(2)) or do not (subsection (c)(1)) take precautions, there is no liability under either state or federal law. As precautions are costly, not only in direct outlay but also in lost revenue from the filtered customers, ISPs may be expected to take the do nothing option and enjoy immunity under § 230(c)(1). Yet § 230(c)—which is, recall, part of the “Communications Decency Act”—bears the title “Protection for ‘Good Samaritan’ blocking and screening of offensive material”, hardly an apt description if its principal effect is to induce ISPs to do nothing about the distribution of

---

responsibilities, be required to monitor all of the content created by third parties *and thereby become derivatively liable for the illegal acts of others.*” *Id.* at 12 (emphasis added). It is certainly possible to interpret the language of Section 230 as providing for such a limitation on liability without disavowing *any* theory of liability that derives in any way from the illegal acts of others.

<sup>47</sup> *Chicago Lawyers’ Committee for Civil Rights Under Law v. Craigslist*, 519 F.3d 666, 671 (7th Cir. 2008) (emphasis added, except where indicated).

<sup>48</sup> *Id.* at 669.

indecent and offensive materials via their services. *Why should a law designed to eliminate ISPs' liability to the creators of offensive material end up defeating claims by the victims of tortious or criminal conduct?*<sup>49</sup>

And as a final Easterbrook opinion rightly points out, this limitation on the scope of immunity is surely tied to the distinction between speech and non-speech causes of action implied by the use of the communication-tort-specific term of art, “publisher”:

Section 230's title, “Protection for private blocking and screening of offensive material”, does not suggest that it limits taxes that have nothing to do with the content of any speech (the City's tax is the same whether the theater is performing “South Pacific” or “Hair”)....

...As earlier decisions in this circuit establish, subsection (c)(1) does not create an “immunity” of any kind. It limits who may be called the publisher of information that appears online. That might matter to liability for defamation, obscenity, or copyright infringement. But Chicago's amusement tax does not depend on who “publishes” any information or is a “speaker”. Section 230(c) is irrelevant.<sup>50</sup>

As the scope of conduct deemed protected by Section 230 has expanded, the defense of Section 230 immunity as based on the moderator's dilemma has grown concomitantly to include both substantive and procedural protection for speech beyond what First Amendment doctrine provides.<sup>51</sup> This growth has occurred notwithstanding that exceptions to the First Amendment exist for “low-value” speech where the harms of protecting the speech are deemed to outweigh the benefits.<sup>52</sup>

The First Amendment does not prevent the suppression of illegal content like child pornography, nor does it protect fraud, perjury, true threats, incitement to violence, and the like.<sup>53</sup> Even with respect to the First Amendment's interaction with defamation law, the Court has held that speech protections are not “justified solely by reference to the interest of the press and broadcast media in immunity from liability,”<sup>54</sup> and the proper balancing between speech and its regulation

---

<sup>49</sup> Doe v. GTE Corp., 347 F.3d 655, 660 (7th Cir. 2003) (emphasis added).

<sup>50</sup> City of Chicago, Ill. v. Stubhub, 624 F.3d 363, 365-66 (7th Cir. 2010). See also Doe, *id.* (“There is yet another possibility: perhaps § 230(c)(1) forecloses any liability that depends on deeming the ISP a ‘publisher’—defamation law would be a good example of such liability—while permitting the states to regulate ISPs in their capacity as intermediaries.”).

<sup>51</sup> See Goldman, *Why Section 230 Is Better*, *supra* note 13, at 36-44.

<sup>52</sup> Child pornography, for instance, is technically speech, but it is not (and should not be) *protected* speech. See, e.g., Osborne v. Ohio, 495 U.S. 103 (1990) (upholding state law outlawing possession and viewing of child porn). Of course, one of the exceptions to Section 230 immunity is for content that is illegal under federal criminal law—an exception adopted with child pornography specifically in mind. See 47 U.S.C. § 230(e)(1). Notably, such exceptions can and do exist to Section 230 as well: Section 230 does not act as a bar to the application of federal criminal law (or state laws consistent with Section 230). Thus, even under the current regime, Congress can adopt a piecemeal approach and pass criminal laws that attempt to solve incentive problem we note throughout this paper.

<sup>53</sup> See Sperry, *Committee Prepares to Grill Tech CEOs*, *supra* note 1.

<sup>54</sup> Gertz v. Robert Welch, Inc., 418 U.S. 323, 343 (1974).

under the First Amendment is not indifferent to the content of the speech at issue: “Like every other case in which this Court has found constitutional limits to state defamation laws, *Gertz* involved expression on a matter of undoubted public concern.”<sup>55</sup> Extending immunity to online intermediaries to remove their responsibility for “low-value” third-party speech may, in a superficial sense, “enhance” speech, but it also allows negative externalities to be imposed upon others.

The same goes for illegal or tortious online conduct that may *appear* to be speech but is not, in fact, solely or primarily speech.<sup>56</sup> It is only by ignoring the costs of such conduct that commentators are able to categorically deem Section 230 “better” than the First Amendment. As Citron & Wittes rightly point out, however, the absence of legal responsibility for online harms engenders a considerable amount of illegal and tortious conduct in addition to the beneficial:

Although § 230 has secured breathing space for the development of online services and countless opportunities to work, speak, and engage with others, it has also produced unjust results. An overbroad reading of the CDA has given online platforms a free pass to ignore illegal activities, to deliberately repost illegal material, and to solicit unlawful activities while ensuring that abusers cannot be identified. Companies have too limited an incentive to insist on lawful conduct on their services beyond the narrow scope of their terms of service. They have no duty of care to respond to users or larger societal goals. They have no accountability for destructive uses of their services, even when they encourage those uses. In addition, platforms have invoked § 230 in an effort to immunize many activities that have very little to do with speech.<sup>57</sup>

The view Citron and Wittes express (“companies... have no accountability for destructive uses of their services”) is surely overstated—after all, the actions of most large tech platforms are subject to an incredible degree of public scrutiny, from congressional hearings to citizen and journalist investigations to public lambasting on those very platforms. Moreover, there are market incentives to remove or reduce the visibility of disfavored speech, including spam, harassing speech, and other types of speech for which most users do not want to be an audience. Profit-maximizing

---

<sup>55</sup> *Dun & Bradstreet, Inc. v. Greenmoss Builders, Inc.*, 472 U.S. 749, 756 (1985).

<sup>56</sup> See, e.g., Danielle Keats Citron & Mary Anne Franks, *supra* note 37, at 61 (“There is no justification for treating the internet as a magical speech conversion machine: if the conduct would not be speech protected by the First Amendment if it occurs offline, it should not be transformed into speech merely because it occurs online.”); see also *id.*, at 56-61 (discussing this theme in detail). For example, soliciting illegal goods technically involved speech, but it is really the conduct of illegal solicitation that is the focus of a legal investigation. Similarly, stalking and harassment may involve speech over the Internet, but it is the conduct of stalking or harassment that is targeted.

<sup>57</sup> Citron & Wittes, *supra* note 46, at 413.



platforms must keep users engaged in order to make money from ad sales, and thus they need to have moderation practices to weed out speech that deters user engagement.<sup>58</sup>

Nonetheless, the sentiments expressed by Citron and Wittes are directionally correct insofar as they identify that an expansive interpretation of Section 230 has limited the application of *legal* incentives to online intermediaries in ways that would influence their behavior. While the government is extremely limited in its ability to shape market demand or change the underlying preferences of platform users, it can deter platforms' facilitation of unlawful conduct by allowing intermediaries to be held legally accountable under certain circumstances. Moreover, absent legal incentives, media scrutiny and market pressures will sometimes be insufficient to deter platforms from soliciting, hosting, or encouraging unlawful content.<sup>59</sup> It is particularly noteworthy that, where Section 230 has not fully barred suit against online intermediaries,<sup>60</sup> "in more than half of the cases studied, plaintiffs succeeded in getting the offensive content removed from the defendants' Web sites or online services."<sup>61</sup> Even though a majority of these cases were ultimately dismissed either on Section 230 or other grounds, it appears that, in at least some cases, only litigation was able to provide recourse for injured plaintiffs; voluntary moderation alone was not.<sup>62</sup> Market and reputational forces are certainly important and significant. But there is every reason to expect that their constraints won't align perfectly with specific policy objectives embodied in the law. At some times, for some people, and with respect to some laws, that may well be a *good* thing. But from the perspective of the implementation of the societal preferences reflected in the law, they are necessarily imperfect.

And, overbroad or not, Citron and Wittes are correct to note that the failure to distinguish between speech and conduct in the application of Section 230 is an important source of the limited application of legal rules to online intermediaries. Just because conduct occurs on an online platform shouldn't automatically make it speech impervious to regulation.

First Amendment doctrine draws a line, contested though it might be, not only between protected and unprotected speech but between speech and conduct.... Because so much online

---

<sup>58</sup> See Ben Sperry, *An L&E Defense of the First Amendment's Protection of Private Ordering*, TRUTH ON THE MARKET (Apr. 23, 2021), <https://truthonthemarket.com/2021/04/23/an-le-defense-of-the-first-amendments-protection-of-private-ordering/>. But see Citron & Franks, *supra* note 37, at 52-53 ("Market forces alone are unlikely to encourage responsible content moderation... [and] keeping up destructive content may make the most sense for a company's bottom line.").

<sup>59</sup> See, e.g., *Jones v. Dirty World Entertainment Recordings LLC*, 755 F.3d 398 (6th Cir. 2014) (in which Nik Richie of TheDirty.com solicited anonymous gossip from users, added editorial notes, and signed it, but the site still received Section 230 immunity in a defamation suit).

<sup>60</sup> See *infra* notes 108 to 119 and accompanying text.

<sup>61</sup> David S. Ardia, *Free Speech Savior or Shield for Scoundrels: An Empirical Study of Intermediary Immunity Under Section 230 of the Communications Decency Act*, 43 LOYOLA LA L. REV. 373, 493 (2010).

<sup>62</sup> See *id.* ("A large proportion of plaintiffs were able to identify and sue the original source of the content that caused them harm. Although their success rate in those suits was quite low, it was not out of line with findings from other studies examining defamation litigation.").

activity involves elements that are not unambiguously speech-related, whether such activities are in fact speech should be a subject of express inquiry. The Court has made clear that conduct is not automatically protected simply because it involves language in some way: “it has never been deemed an abridgement of freedom of speech or press to make a course of conduct illegal merely because the conduct was in part initiated, evidenced, or carried out by means of language, either spoken, written, or printed.”<sup>63</sup>

Section 230 elides this distinction by treating everything online as presumptively protected speech, at least from the point of view of holding platforms liable for failing to prevent harms that emerge from them. The result ends up deterring lawsuits on occasions where offline intermediaries would be held accountable for the same foreseeable harms.<sup>64</sup>

For instance, at a wine festival where wineries can sell their wares to attendees, both the seller and the festival itself could be held liable if a seller failed to check IDs and an underage attendee bought its wine and got into an accident after drinking it.<sup>65</sup> But if *onlinewinesales.com* (not real) set up a platform where wineries could sell wine online,<sup>66</sup> it could theoretically hide behind Section 230 immunity if one of the wineries on its platform similarly sold wine without adequately checking purchasers’ IDs.<sup>67</sup> Indeed, this is essentially what happened in one case where an online platform for private gun sales escaped liability using a Section 230 defense. One of the sellers on its platform failed to perform a legally required background check, and it led to the murder of a woman by her husband against whom she had a restraining order.<sup>68</sup> Put differently, speech maximization is only as

---

<sup>63</sup> Citron & Franks, *supra* note 37, at 58-59 (quoting *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490, 502 (1949)). See also *United States v. Ulbricht*, 31 F. Supp. 3d 540, 568 (S.D.N.Y. 2014) (“[T]he law is clear that speech which is part of a crime is not somehow immunized. For instance, no one would doubt that a bank robber’s statement to a teller—‘This is a stick up’—is not protected speech.”).

<sup>64</sup> See Part III.C, *infra*, for examples of how offline intermediaries are held liable for harms in particular situations under the common law.

<sup>65</sup> See discussion in Part III.C on dram shop liability. See also Katherine Kokal, *2 Lawsuits Filed After Punches Thrown at Sea Pines Festival. Here’s Where They Stand*, THE ISLAND PACKET (May 2, 2019), <https://www.islandpacket.com/article229900489.html> (“Sea Pines Resort and Hilton Head real estate partner Tad Segars have settled a lawsuit in which Segars says he was punched in the face by an intoxicated attendee at the 2016 Hilton Head Wine & Food Festival. Segars sued the wine and food festival, the resort and Coastal Security Services for negligence and dram shop law liability. That case was settled on March 5 for an undisclosed amount, according to court filings in the Beaufort County Court Index.”).

<sup>66</sup> Obviously other laws could complicate this possibility for the sellers, which are not the platform itself in this hypo.

<sup>67</sup> This example is not as outlandish as one might think. One study from 2012 found that of 100 orders placed by underage buyers online, 45% were successfully received and 28% were rejected as a result of age verification. Most vendors (59%) used only weak age verification. Of the successful underage orders, 51% used no age verification at all. See Rebecca S. Williams & Kurt M. Ribisl, *Internet Alcohol Sales to Minors*, 166 ARCH PEDIATR. ADOLESC. MED. 808 (2012), available at <https://jamanetwork.com/journals/jamapediatrics/fullarticle/1149402>.

<sup>68</sup> See *Daniel v. Armslist, LLC*, 926 N.W.2d 710 (Wis. 2019). This result has been defended on speech grounds. See Cathy Gellis, *The Wisconsin Supreme Court Gets Section 230 Right*, TECHDIRT (May 1, 2019), <https://www.techdirt.com/articles/20190501/07150142120/wisconsin-supreme-court-gets-section-230-right.shtml> (“[A]s we

beneficial as the types of speech that are maximized. Reforms that aim to reduce low-value, third-party speech (which may be conduct disguised, by its online nature, as speech) by using the threat of liability to encourage different or more aggressive moderation practices could *increase* social welfare if the benefits of those reductions outweigh the costs of the lost speech.

This is not to say that there are no valid reasons to limit the application of such liability in some online contexts. But it highlights that broad application of Section 230 can prevent the operation of the sort of intermediary-liability laws that are commonly employed in offline contexts to increase the effectiveness of direct enforcement and further limit the incidence of harmful conduct in the first place. At the very least, it would be difficult to maintain in such circumstances, as Chris Cox has asserted, that “Section 230 operates to ensure that like activities are always treated alike under the law.... Whether in the offline world or the internet, the same legal rules and responsibilities apply across the board to all”<sup>69</sup>

## B. “Death by ten thousand duck-bites”

Procedurally, Section 230 immunity protects would-be defendants not just from liability for harm caused by third-party content, but also from having to incur a substantial share of the attendant litigation costs. It does this by facilitating early motions to dismiss before evidentiary discovery.<sup>70</sup> The two costs—liability costs and litigation costs—are related but distinct.

Many argue<sup>71</sup> that holding online intermediaries liable for failing to remove offensive content would lead to a flood of lawsuits that would ultimately overwhelm service providers, and

---

pointed out in our briefs, there is always more at stake than just the case at hand. Whittling away at Section 230’s important protection because one plaintiff may be worthy leaves all the other worthy online speech we value vulnerable. It is protected only when platforms are protected. When their protection is compromised, so is all the speech they carry. Which is why it is so important for courts to resist the emotion stirred by instant facts and clinically apply the law as it was written, so that instead of helping just one person it will help everyone.”).

<sup>69</sup> Cox, *PACT Act Testimony*, *supra* note 35, at 12.

<sup>70</sup> See Goldman, *Why Section 230 Is Better*, *supra* note 13, at 39-44. See also *Nemet Chevrolet, Ltd. v. ConsumerAffairs.com, Inc.*, 591 F.3d 250, 255 (4th Cir. 2009) (“We thus aim to resolve the question of § 230 immunity at the earliest possible stage of the case because that immunity protects websites not only from ‘ultimate liability,’ but also from ‘having to fight costly and protracted legal battles.’”).

<sup>71</sup> See, e.g., Berin Szóka, Ashken Kazaryan, & Jess Miers, *Why Section 230 Matters and How Not To Break the Internet; DOJ 230 Workshop Review, Part I*, TECHDIRT (Feb. 21, 2020), <https://www.techdirt.com/articles/20200221/11290843961/why-section-230-matters-how-not-to-break-internet-doj-230-workshop-review-part-i.shtml> (“One duck-bite can’t kill you, but ten thousand might. Likewise, a single lawsuit may be no big deal, at least for large companies, but the scale of content on today’s social media is so vast that, without Section 230, a large website might face far more than ten thousand suits. Conversely, litigation is so expensive that even one lawsuit could well force a small site to give up on hosting user content altogether. A single lawsuit can mean death by ten thousand duck-bites: an extended process of appearances, motions, discovery, and, ultimately, either trial or settlement that can be ruinously expensive. The most cumbersome, expensive, and invasive part may be “discovery”: if the plaintiff’s case turns on a question of fact, they can force the defendant to produce that evidence. That can mean turning a business inside out – and protracted fights over what evidence you do and don’t

sub-optimally diminish the value these firms provide to society—a so-called “death by ten thousand duck-bites.”<sup>72</sup> Relatedly, firms that face potentially greater liability would be forced to internalize some increased—possibly exorbitant—degree of compliance costs even if litigation never materialized.

Concern for judicial economy and operational efficiency are laudable, of course, but such concerns are properly addressed toward minimizing the costs of litigation in ways that do not undermine the deterrent and compensatory effects of meritorious causes of action.<sup>73</sup> Litigation costs that exceed the minimum required to properly assign liability are deadweight losses to be avoided;<sup>74</sup> liability costs—when properly found—are not a deadweight cost.<sup>75</sup> They ought to be borne by the party best positioned to prevent harm.

Litigation costs can therefore be further separated (broadly) into two broad types: 1) litigation costs necessary to properly find liability, and 2) “unnecessary” litigation costs, including both litigation costs incurred in the course of unmeritorious litigation and any “avoidable” litigation costs incurred in the course of meritorious litigation that exceed the minimum required.

These three costs—liability costs, necessary litigation costs, and excess litigation costs—are regularly conflated. Discussions of the litigation effects of Section 230 often count the avoidance of liability costs as a benefit, for example. If your concern is to avoid overly burdening online intermediaries with costs, it may not matter if the source of the cost is a legitimate liability award or

---

have to produce. The process can easily be weaponized, especially by someone with a political ax to grind.”). See also, e.g., Goldman, *Why Section 230 Is Better*, *supra* note 13; Mike Masnick, *Hello! You’ve Been Referred Here Because You’re Wrong About Section 230 of the Communications Decency Act*, TECHDIRT (Jun. 23, 2020), <https://www.techdirt.com/articles/20200531/23325444617/hello-youve-been-referred-here-because-youre-wrong-about-section-230-communications-decency-act.shtml>.

<sup>72</sup> *Roommates.com*, 521 F. 3d at 1174 (“Websites are complicated enterprises, and there will always be close cases where a clever lawyer could argue that *something* the website operator did encouraged the illegality. Such close cases, we believe, must be resolved in favor of immunity, lest we cut the heart out of section 230 by forcing websites to face death by ten thousand duck-bites, fighting off claims that they promoted or encouraged—or at least tacitly assented to—the illegality of third parties.”).

<sup>73</sup> See, e.g., Jonathan L. Molot, *How U.S. Procedure Skews Tort Law Incentives*, 73 INDIANA L.J. 59, 60 (1997) (“Corporations, politicians, and the media generally share the sense that litigation in the United States is inordinately expensive and that our system of litigation thus deters productive conduct.... In particular, lawyers and clients alike understand that the cost of litigation may affect outcomes. It is less obvious, however, which procedural rules contribute to the costliness of litigation and whether these rules together lead more often to plaintiffs foregoing meritorious suits, to defendants paying for meritless ones, or to parties settling meritorious suits early and thereby avoiding the costs of litigation entirely. Without further inquiry, it is impossible to determine whether expensive litigation leads defendants to expect to pay more, less, or the same amounts for suits as they would under substantive law alone. Such expectations are at the core of any analysis of how procedural rules affect deterrence.”).

<sup>74</sup> Or, more accurately, they *may* be. Litigation costs provide some of the deterrent effect of liability rules, and optimization of those rules properly accounts not only for liability awards but also litigation costs. To the extent that “excess” litigation costs are necessary for optimal deterrence they are also not a “deadweight cost to be avoided.”

<sup>75</sup> For a more comprehensive analysis of these points, see Section III.B, *infra*.

an excess litigation expense. Indeed, Section 230 offers freedom from *liability* for online publishers of third-party content. But is liability really the problem, or is it (excess) litigation costs?

Unsurprisingly, defenses of Section 230 immunity usually cite litigation costs explicitly, not liability risks. But the latter is usually swept into the former. Emblematic is Eric Goldman's discussion of the issue:

Section 230(c)(1)'s early dismissals are valuable to defendants. They reduce the defendant's out-of-pocket costs to defeat an unmeritorious claim. For smaller Internet services, defending a single protracted lawsuit may be financially ruinous. Also, complex litigation can divert substantial managerial and organizational attention and mindshare from maintaining or enhancing the service. Thus, the ability of a defendant to resolve a case on a motion to dismiss (and avoiding expensive discovery) protects small and low-revenue Internet services, which in turn enhances the richness and diversity of the Internet ecosystem.<sup>76</sup>

The explicit reference to litigation costs obscures the point about liability risk. Goldman's first claim is limited, in his telling, to unmeritorious cases, but not so his subsequent assertions: "Protracted lawsuits" and "complex litigation" may well be the *necessary* costs of optimal liability. The claims of benefits here are indifferent to the merits of the litigation. But presumably *plaintiffs* are not indifferent. *Victims of tortious conduct* are not indifferent. *Society* may not be indifferent.

We should all welcome efforts to reduce *unnecessary* litigation costs (no matter the context). But we should be skeptical of policy proposals that simply ignore the countervailing costs of precluding *meritorious* lawsuits.

In his 2010 study of litigation involving Section 230 claims, David Ardia concludes that at least some types of providers "would likely fare far worse under the common law"<sup>77</sup> that preceded Section 230, and that would have continued to develop were it not for the interposition of Section 230.<sup>78</sup> But Ardia notes that the reason they would fare worse is that, in many cases, service providers would have been identified as relevant distributors or publishers capable of controlling the

---

<sup>76</sup> Goldman, *Why Section 230 Is Better*, *supra* note 13, at 40-41.

<sup>77</sup> Ardia, *supra* note 61, at 479.

<sup>78</sup> *Id.* at 480 ("Would the common law have evolved some way to grant these intermediaries a presumption of non-liability akin to the common law's approach to conduit liability? We simply do not know.").

proliferation of tortious content.<sup>79</sup> Thus, far from being indifferent, the benefits the service providers enjoy are at *the expense* of plaintiffs who would otherwise have been able to seek legal recourse.<sup>80</sup>

Defenses of Section 230 that invoke litigation costs often fail to consider the social costs that Section 230 immunity might engender—that is, the lost benefits that would accrue to plaintiffs with meritorious cases that are never brought because the statute effectively prohibits them. Thus, it is quite possible that, in avoiding the duck-bites problem, Section 230 immunity has imposed social costs greater than the marginal benefits the immunity provides in many classes of cases.

As Judge Kozinski suggested in his *Roommates.com* opinion, there are at least three types of cases that might give rise to litigation-defense costs: close cases, frivolous cases, and meritorious cases. The court in *Roommates.com* rightly focused on the close cases—cases “where a clever lawyer could argue that *something* the website operator did encouraged the illegality.”<sup>81</sup> Such cases are relevant for two reasons. First, they may or may not be meritorious, and the incentives to bring such cases—which depend, in part, on the prospects that the plaintiff will prevail—may shift with even fairly small adjustments to the rules of civil procedure. There is good reason to be concerned about the difficulty in setting these complex rules optimally. Erring on the side of immunity may, therefore, be appropriate, lest anything short of immunity over-encourage vexatious litigation (and thus “death by ten thousand duck-bites”).<sup>82</sup> Second, these cases are relevant because, all else equal, they should lead to only marginal benefits, even when meritorious and successful. That is, the close cases are likely cases where a platform’s conduct was only marginally problematic, or where the evidence of negligence or harm are only barely persuasive. Winning these cases may offer only minimal social benefits relative to the cost of litigation, because they would deter only marginally problematic conduct.

---

<sup>79</sup> *Id.* (“For these intermediaries, the editorial-control distinction in the common law breaks down because virtually all content hosts and search/application providers have the ability to exercise editorial control over third-party speech, even if they do not choose to exercise that power. Indeed, they would likely fare far worse under the common law than predicted if courts were to apply publisher liability to an intermediary that simply has the ability to exercise editorial control.”).

<sup>80</sup> Notably, in Ardia’s analysis most providers in his study would likely not have been found liable even under the common law—just the subset that enjoyed the immunity despite what would have been otherwise possible. *See id.* at 480 (“[M]any of the intermediaries that invoked section 230 likely would not have faced eventual liability under the common law because they lacked knowledge of and editorial control over the third-party content at issue in the cases.”).

<sup>81</sup> *Roommates.com*, 521 F. 3d at 1174.

<sup>82</sup> Importantly, however, even the *Roommates.com* court recognizes that this notion has its limits, particularly when it comes to the scope of the rule:

However, a larger point remains about the scope of immunity provisions. It’s no surprise that defendants want to extend immunity as broadly as possible. We have long dealt with immunity in different, and arguably far more important, contexts—such as qualified immunity for police officers in the line of duty—and observed many defendants argue that the risk of getting a close case wrong is a justification for broader immunity. Accepting such an argument would inevitably lead to an endless broadening of immunity, as every new holding creates its own borderline cases.

*Roommates.com*, 521 F.3d at 1175 n.39 (citation omitted).

But presumably no one thinks we should deter clearly meritorious cases, nor encourage clearly frivolous ones. The problem with an absolute immunity regime, however, is that it fails to distinguish among types of cases, and necessarily precludes meritorious suits, even as it prevents frivolous ones and close cases. And, of course, it is the clearly meritorious suits that, on the margin, would yield the most beneficial changes in conduct and/or the most compensation for the most egregious harms. Without consideration of the cost of precluding meritorious litigation, it is impossible to evaluate the propriety of blanket immunity.

A crucial problem is that we lack empirical data on the counterfactual: what the mix of cases and the costs of defending some or all of the underlying claims currently immunized by Section 230 would be were the law to be repealed or greatly circumscribed.<sup>83</sup> Opponents of Section 230 reform assert that the consequences would be catastrophic, especially for smaller platforms.<sup>84</sup> But while expanded liability surely would drive increased litigation costs to some extent, it is far from certain by how much, with what effect on platform and user activities, and what relative burdens it would place on smaller or larger intermediaries.

To begin, it must be noted that economic actors face liability risk all the time in many contexts, even as intermediaries.<sup>85</sup> The mere fact of increased liability risk is not, in itself, a persuasive objection to changes in legal standards; it could very well be that the current liability risk is *too low*.<sup>86</sup> Indeed, neither the law nor common sense necessarily supports the arguments for blanket immunity that Section 230's proponents often make:

We could, of course, decide that Internet intermediaries should never be liable for the misconduct of others.... This is not the rule in tort law generally, however. Before we confer such a broad immunity on Internet intermediaries, we need to consider whether this is what we really intend. For example, do we really intend to absolve Internet intermediaries of any liability for failing to take reasonable measures to deter credit card fraud? Such fraud is, after all, misconduct by others, and fraudulent credit card numbers seem to be “information provided by another information content provider.” As described above, the rationale of

---

<sup>83</sup> The closest to quantification on these costs is an estimate from one advocacy group based upon self-reported estimates from lawyers of how much litigation would be expected to cost. See Evan Engstrom, *Primer: Value of Section 230*, ENGINE (Jan. 31, 2019), <https://www.engine.is/news/primer/section230costs>.

<sup>84</sup> See, e.g., Eric Goldman, *Want to Kill Facebook and Google? Preserving Section 230 Is Your Best Hope*, BALKANIZATION (Jun. 3, 2019), <https://balkin.blogspot.com/2019/06/want-to-kill-facebook-and-google.html>.

<sup>85</sup> See *infra* Part III.C for a discussion of several of these contexts.

<sup>86</sup> See, e.g., Danielle Citron & Quinta Jurecic, *Platform Justice: Content Moderation at an Inflection Point*, Hoover Working Group on National Security, Technology, and Law, Aegis Series Paper No. 1811 (Sept. 5, 2018) at 4, available at [https://www.hoover.org/sites/default/files/research/docs/citron-jurecic\\_webreadypdf.pdf](https://www.hoover.org/sites/default/files/research/docs/citron-jurecic_webreadypdf.pdf) (“Although Section 230 has secured breathing space for the development of online services and countless opportunities to work, speak, and engage with others, it has also given platforms a free pass to ignore destructive activities, to deliberately repost illegal material, and to solicit unlawful activities while ensuring that abusers cannot be identified.”).

avoiding collateral censorship stops well short of such broad immunity; if we want greater immunity, some other justification is required.<sup>87</sup>

Moreover, as noted, we cannot anticipate, *ex ante*, what the cost of litigation would be in the absence of immunity. The one published effort to quantify the relevant considerations<sup>88</sup> is of some slight help, but it does nothing to explain the relative weights of the expected costs, which range in the report from \$0 to more than \$500,000.<sup>89</sup> We also have little sense of how likely it would be that any given piece of content would incur any given cost within that enormous range. Nor, of course, do we know the expected *returns* from any given item or type of content: How often will it be the case that expected litigation costs outweigh the expected benefits of allowing certain user-generated content?

A further background presumption to this discussion is that greater liability risk will give rise to many more non-meritorious suits. This distinction is often left unclear in writings critical of Section 230 reform, but it is extremely important. There is a big difference between assuming that *any* additional expected litigation cost is unwarranted, regardless of the merits, and assuming that the increased likelihood and cost of *unmeritorious or vexatious* litigation will necessarily outweigh the benefits of legitimate litigation. To the extent that the concern is with the latter (as we believe it should be), it is far from certain that removing all possibility of liability is the only or best way to control the costs of vexatious litigation.

It is also important to consider what “death” in the “death by ten thousand duck-bites” metaphor entails. When opponents of Section 230 reform worry about the cost of litigation in the absence of the law’s grant of immunity, they appear to assume that the current amount of user-generated content is optimal and must remain constant.<sup>90</sup> But the scale of user-generated content *need not be* so vast, nor websites so large. For any given platform, the extent of litigation risk probably declines with the amount of content, all else equal. Thus, in the face of liability, any website likely could reduce its risk of liability simply by hosting less content.<sup>91</sup> (Of course, that isn’t all it can do: it can also screen users and/or content to ensure that potentially illegal content is less likely to get through). That may be a cost of a liability regime, but it isn’t quite “death.”

---

<sup>87</sup> Wu, *supra* note 11, at 341.

<sup>88</sup> See Engstrom, *supra* note 83.

<sup>89</sup> *Id.*

<sup>90</sup> See, e.g., Szóka, et al, *supra* note 71 (“[T]he scale of content on today’s social media is so vast that, without Section 230, a large website might face far more than ten thousand suits.”).

<sup>91</sup> This could take many forms, and it’s unlikely to be as simplistic as numerically reducing the amount of raw content hosted. Instead, through some set of filtering mechanisms that more precisely select users, content types, and posting frequency, platforms would adjust their scale in proportion to their ability to moderate. The net effect could be a lower volume of content per platform (and possibly a larger number of specialized platforms).



Relatedly, some argue that Section 230 immunity is pro-competitive. That is, the claim is that reform to the current immunity regime could create barriers to entry for startups, as larger online intermediaries can afford to dedicate resources to moderating content and defending lawsuits that smaller online intermediaries simply cannot.<sup>92</sup> Eric Goldman, for example, has claimed that:

Those *disruptive innovators absolutely require legal immunity* to grow big and popular enough to change consumer practices and gain consumer loyalty, without being swamped by lawsuits and the high costs of content moderation obligations. *Section 230 is an essential piece to ensure that future Google- and Facebook-killers have a chance of emerging.*<sup>93</sup>

But this claim needs to be counterposed with the possibility that Section 230 immunity subsidizes *diseconomies* of scale, as well. If moderation at scale is as difficult as many argue,<sup>94</sup> and otherwise meritorious lawsuits are deterred by Section 230, then it could be the case that some platforms are larger than they would have been absent Section 230 immunity. In other words, it isn't clear, on balance, whether Section 230 immunity has increased or decreased concentration among online intermediaries, or that the current scale of platforms is optimal, all things considered.

Indeed, given that, as discussed above, one way to mitigate liability risk and litigation cost under a regime without Section 230 immunity would be to reduce the number of users or the amount of content they post, it is no coincidence that current antitrust arguments against Big Tech companies are regularly intertwined with Section 230 arguments; the two go hand-in-hand.<sup>95</sup> If your concern is that Facebook, for example, is “too big,” the argument that weakening Section 230 could

---

<sup>92</sup> See, e.g., Masnick, *supra* note 71; Tim Wu, *Why Both Liberals and Conservatives Are Completely Wrong About Section 230*, PROMARKET (Dec. 13, 2020), <https://promarket.org/2020/12/13/liberals-conservatives-wrong-section-230-reform-repeal/> (“Abolishing Section 230 would not address disinformation and propaganda on social media nor charges of anti-conservative censorship. But its repeal would probably hurt startups and smaller rivals, further insulating big platforms from competition.”); Amrita Khalid, *Why Startups Have So Much Riding on Section 230’s Future*, INC. (Jan. 15, 2021), <https://www.inc.com/amrita-khalid/section-230-communications-decency-act-tech-startups-donald-trump.html>; Ryan Nabil, *Why Repealing Section 230 Will Hurt Startups and Medium-Sized Online Businesses*, COMPETITIVE ENTERPRISE INSTITUTE (Feb. 1, 2021), <https://cei.org/blog/why-repealing-section-230-will-hurt-startups-and-medium-sized-online-businesses/>.

<sup>93</sup> Goldman, *Want to Kill Facebook and Google?*, *supra* note 84 (emphasis added).

<sup>94</sup> See, e.g., Tarleton Gillespie, *Platforms Are Not Intermediaries*, 2 GEO. L. TECH. REV. 198, 198 (2018) (“Content moderation is such a complex and laborious undertaking, it is amazing that it works at all and as well as it does. Moderation is hard. This should be obvious, but it is easily forgotten. Policing a major platform turns out to be a resource intensive and relentless undertaking; it requires making difficult and often untenable distinctions between the acceptable and the unacceptable; it is wholly unclear what the standards for moderation should be, especially on a global scale; and one failure can incur enough public outrage to overshadow a million quiet successes.”).

<sup>95</sup> See, e.g., *Proposals to Strengthen the Antitrust Laws and Restore Competition Online: Hearing Before the Subcomm. On Antitrust, Commercial and Admin. Law* (Oct. 1, 2020) (Statement of Rachel Bovard), *available at* <https://docs.house.gov/meetings/JU/JU05/20201001/111072/HHRG-116-JU05-Wstate-BovardR-20201001.pdf> (arguing for both Section 230 reform and antitrust enforcement against Big Tech); Makena Kelly, *Sen. Josh Hawley is Making the Conservative Case Against Facebook*, THE VERGE (Mar. 19, 2019), <https://www.theverge.com/2019/3/19/18271487/josh-hawley-senator-missouri-republican-facebook-google-antitrust-data-privacy> (noting that Sen. Hawley favors both Section 230 reform and antitrust enforcement against Big Tech for anti-conservative bias, among other things).

lead to huge litigation risk and considerably less content aren't problems to contend with; they're part of the solution.<sup>96</sup>

While there are ubiquitous claims regarding the dire consequences of increased liability stemming from Section 230 reform, the relationship between platform size (number of users and/or amount of user-generated content) and litigation cost is, for all intents and purposes, completely unexplored. It may be largely monotonic—that is, the expected cost of litigation may increase in a constant ratio with the increase in platform size—but it is unlikely to be monotonic over the entire size range. As many Section 230 defenders argue, for example, it is surely the case that, for some platforms, even a small degree of litigation risk could lead to *no* content (i.e., the absence of the platform altogether), not just *less* content. By the same token, perhaps only a small fraction of users and user content presents any appreciable litigation risk, and a relatively small reduction in users who post “*bad*” content could enable a *substantial* reduction in liability risk and expected litigation cost. Assessing the actual likely consequences of reduced immunity thus entails understanding how well platforms can estimate expected liability risk from any given user, group of users, or type of content, and how effectively they can manage litigation risk by targeting a smaller number of particularly problematic users or types of content.

This presents the important possibility that an increase in liability risk may lead not to substantial increases in litigation costs, but to other changes that may be less privately costly to a platform than litigation, and which may be socially desirable. Among these changes may be an increase in preemptive moderation (“collateral censorship”); smaller, more specialized platforms and/or tighter screening of platform participants on the front end (both of which are likely to entail stronger reputational and normative constraints); the establishment of more effective user-reporting and harm-mitigation mechanisms; development and adoption of specialized insurance offerings; or any number of other possible changes.

---

<sup>96</sup> It should be noted (although it never is) that the reverse dynamic could occur for platform *users*. Perhaps, in the face of more moderation because of liability risk, larger, more established third-party users could become better entrenched against their own smaller competitors or new entrants. Consider an advertising platform, for example. A liability regime could impose upon the platform operator the obligation to review submitted ads for harmful content, adding delay, cost, and possibly restrictions on even non-harmful content, all of which would impose some costs or reduce quality on the advertiser. But what if the advertiser is a large, established company? Undoubtedly the large third-party advertiser would be in a better position to negotiate rates, placement, and expedited moderation approval with the platform operator as compared to smaller rivals. Certainly the risk of illegal content originating from a large established company would be lower, thus the platform would have some incentive to be relatively more permissive. A start-up competitor would be unlikely to receive a similar dispensation, nor should it as its behavior increases the liability risk to the platform operator out of proportion to its ad-spend value to the platform (relative to the larger competitor). Thus, a regime that introduces greater liability for platforms could have deleterious downstream effects on competition among platform users. See Benjamin Edelman, *Least-Cost Avoiders in Online Fraud and Abuse*, IEEE SECURITY & PRIVACY (Jul./Aug. 2010) 78, 80 (noting that reasonable care could entail that “[k]nown-trustworthy advertisers could be exempt from unnecessary delays. Conversely, in light of the deception so prevalent in ‘free’ offers, the provider could flag any advertiser promising free service for heightened review”).

Finally, it must be noted that, if the cost of litigation is a problem, it is a problem throughout the economy, not just one faced by online platforms. Millions of small businesses confront expected litigation costs, and large businesses like Walmart, for example, are sued literally every day (in Walmart's case, almost certainly multiple times daily). It is unclear what separates this litigation risk from the risk facing online platforms. Most would answer, perhaps, that the risk for online platforms is indirect; it is a function of bad acts by third parties, not actions by the platforms themselves. But a litigation cost is a litigation cost; why should the specific type of litigation matter if companies are deterred from forming or from expanding their activity because of excessive litigation risk? Moreover, it is not clear that the rate of *non-meritorious* lawsuits is any higher in the third-party/indirect liability or online contexts. And it is only non-meritorious lawsuits we should wish to deter. Counting the cost of defending meritorious lawsuits as an avoidable and unfortunate expense is tantamount to wishing away our civil-justice system. That is unlikely to be a defensible position in any regard, but it certainly is not defensible solely in the context of online platforms.

There are, of course, some differences between the online and offline worlds, but they are primarily differences of degree, not of kind.<sup>97</sup> It is certainly the case, for example, that moderation at scale is essentially the norm online and a difficult problem for online platforms of virtually all sizes to tackle.<sup>98</sup> Given the potential scale of even the smallest online services, moderating millions or billions of pieces of content will inevitably lead to some bad content slipping through algorithmic filters and user reporting. Thus, to some greater or lesser extent, overly broad reforms to Section 230—those that introduce liability without adjusting for the fundamental scale problems that plague online services—*could* lead to widespread chilling of a large amount of speech. But the same overbreadth problem, in reverse, plagues Section 230 immunity in its current incarnation:

In creating the Law of Cyberspace, Congress did the opposite of what Judge Easterbrook had urged: rather than clarifying existing legal principles—in particular, principles of immunity,

---

<sup>97</sup> Moreover, there is a danger in essentializing the “online-ness” of conduct and viewing online harms as a special class of harms, rather than as particular occurrences of harms that could arise anywhere. See Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL FORUM 207, 207-08 (“[T]he best way to learn the law applicable to specialized endeavors is to study general rules. Lots of cases deal with sales of horses; others deal with people kicked by horses; still more deal with the licensing and racing of horses, or with the care veterinarians give to horses, or with prizes at horse shows. Any effort to collect these strands into a course on ‘The Law of the Horse’ is doomed to be shallow and to miss unifying principles.... Only by putting the law of the horse in the context of broader rules about commercial endeavors could one really understand the law about horses.”). *But see* Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501 (1999).

<sup>98</sup> See, e.g., Daisy Soderberg-Rivkin, *Five myths about online content moderation, from a former content moderator*, R STREET (Oct. 30, 2019), <https://www.rstreet.org/2019/10/30/five-myths-about-online-content-moderation-from-a-former-content-moderator/> (“Section 230 can be thought of as giving birth to, and making possible, content moderation. If lawmakers rescind Section 230 protection, tech companies will be open to a lawsuit every time a moderator decides to remove content or leave it on the platform.... Given the astronomical amount of content uploaded to platforms each day... , many companies would likely opt to allow the vilest content to remain on their platforms rather than risking the myriad lawsuits and fines that could easily put them out of business.”). *See also* Jillian C. York & Corynne McSherry, *Content Moderation Is Broken. Let Us Count the Ways*, ELECTRONIC FRONTIER FOUNDATION (Apr. 29, 2019), <https://www EFF.ORG/deeplinks/2019/04/content-moderation-broken-let-us-count-ways>.

complicity, free speech, criminal law, or tort—in light of technological advances and applying those principles to Internet cases, Congress effectively upended all those principles in order to accommodate the supposedly exceptional nature of the Internet.<sup>99</sup>

What is called for is a properly scoped reform that applies the same political, legal, economic, and other social preferences offline as online, aimed at ensuring that we optimally deter illegal content without losing the benefits of widespread user-generated content. Properly considered, there is no novel conflict between promoting the flow of information and protecting against tortious or illegal conduct online. While the specific mechanisms employed to mediate between these two principles online and offline may differ—and, indeed, while technological differences can alter the distribution of costs and benefits in ways that must be accounted for—the fundamental principles that determine the dividing line between actionable and illegal or tortious content offline can and should be respected online, as well. Indeed, even Google has argued for exactly this sort of parity, recently calling on the Canadian government to “take care to ensure that their proposal does not risk creating different legal standards for online and offline environments.”<sup>100</sup>

### III. A Law & Economics Framework for Section 230 Reform

Addressing the objections outlined above and devising viable reforms of Section 230 requires understanding the costs and benefits of moving from the status quo to a new regime. In this, of course, it is the *marginal* that matters. The relevant questions are: To what degree would shifting the legal rules governing platform liability increase litigation costs, increase moderation costs, constrain the provision of products and services, increase “collateral censorship,” and impede startup formation and competition, all *relative to the status quo*, not to some imaginary ideal state? Assessing the marginal changes in all these aspects entails, first, determining how they are affected by the current regime. It then requires identifying both the direction and magnitude of change that would result from reform. Next, it requires evaluating the corresponding *benefits* that legal change would bring in increasing accountability for tortious or criminal conduct online. And finally, it necessitates hazarding a best guess of the *net* effect.

It is unfortunate that the Section 230 reform conversation has been dominated by intuition and assertions about the effects of changing the law, and not by a rigorous calculation of the costs and benefits of such changes. Defenses of the status quo invariably rest on a set of assumptions regarding the direction of change (always negative) and the magnitude of harm (always ill-defined,

---

<sup>99</sup> Mary Anne Franks, *How the Internet Unmakes Law*, 16 OHIO ST. TECH. L.J. 10, 13 (2020) (referencing Easterbrook, *Cyberspace and the Law of the Horse*, *supra* note 97).

<sup>100</sup> Google, *Submission to Canadian Government’s Proposal to Address Online Harms*, Sep. 24, 2021, at 1, available at <https://drive.google.com/file/d/1bvqDtYgyjjagk1YKcKbIhc-m6fu053g/view>. In an accompanying blog post, Google Canada states that “Our starting point is that we believe the same standards should apply to expression in online and offline environments.” See *Our Shared Responsibility: YouTube’s response to the Government’s proposal to address harmful content online*, GOOGLE OFFICIAL CANADA BLOG (Nov. 5, 2021), <https://canada.googleblog.com/2021/11/our-shared-responsibility-youtubes.html>.

but typically catastrophic). Never do these efforts begin from a realistic baseline, consider that there is a wide range of possible outcomes far short of “catastrophic,” or compare (or generally even acknowledge) the corresponding benefits that would be realized.

In fairness, estimating the net effects with any degree of rigor is, indeed, extremely difficult. We do not purport here to offer such an assessment. Instead, what we offer is a high-level discussion of some of the key factors that must be incorporated into any evaluation of possible Section 230 reforms that impose some form of third-party liability on online platforms. Even without knowing these elements with any degree of specificity, the process of considering each—in isolation and in conjunction—serves to undermine breathless claims that Section 230 reform could bring only catastrophic social harm. This alone does not tell us that Section 230 reform is a good idea, but it does give reason for hope. More importantly, it offers crucial guidance on how to shape reform efforts to maximize the prospects that reform will be a net positive.

### **A. Baseline moderation and litigation risks**

Many of the most vehement objections to Section 230 reform are predicated on fundamental misunderstandings of the relevant baseline moderation and liability risks that would obtain in the law’s absence. For example, platforms that face a moderator’s dilemma will have incentives to remove more content than is strictly necessary, so that they might avoid even the potential for liability. But the relevant alternative to a moderation regime in the absence of Section 230 is not “no moderation,” but “self-censorship”: the self-moderation exercised by users themselves in the face of their *own* liability risk. As Felix Wu has observed:

We therefore need to understand what makes collateral censorship a problem. *In particular, the problem cannot be simply that the threat of liability results in the suppression of speech, for that is true whenever there is liability for speech. People regularly engage in self-censorship under fear of liability*, but if that is the crux of the problem, then the appropriate solution would be to change the substantive liability itself.

The unique harm of collateral censorship, as opposed to self-censorship, lies in the incentives that intermediaries have to suppress more speech than would be withheld by original speakers. This additional suppression occurs because intermediaries have different incentives to carry particular content than original speakers have to create it in the first place.<sup>101</sup>

---

<sup>101</sup> Wu, *supra* note 11, at 296-97 (emphasis added). Note, this does not mean that imposing intermediary liability can never lead to over-deterrence; it does mean, however, that the amount of possible over-deterrence is not as large as commonly assumed.

Putting the same point in economic terms, intermediary liability makes sense when a party is in a position “to detect and deter bad acts” or “account for significant negative externalities that are unavoidably associated with its activities.” See Lichtman & Posner, *supra* note 24, at 231-32. But there is a danger in imposing liability when the costs of liability are too high and would “inadvertently interfere with substantial legitimate [] activity.” *Id.* at 232-33. In the case of online speech platforms, the fear is that through monitoring speech or reducing activity level, a substantial amount of legitimate user speech would be harmed. The question, as always, is one of relative costs, including transaction costs, and benefits.

To the extent that users are already exposed to the underlying legal liability from which platforms are immune, the magnitude of “additional censorship” due to intermediary liability will be at least marginally smaller than critics of reform assume. Indeed, if the extent of self-censorship from expanding liability to platforms does prove to be much more significant, this would likely be evidence that users currently believe they can act with impunity on the Internet—further highlighting that Section 230 stymies the normal operation of the law.

A proper evaluation of the merits of an intermediary-liability regime must therefore consider whether user liability alone is insufficient to deter bad actors, either because it is too costly to pursue remedies against users directly,<sup>102</sup> or because the actions of platforms serve to make it less likely that harmful speech or conduct is deterred. The latter concern, in other words, is that intermediaries may—intentionally or not—facilitate harmful speech *that would otherwise be deterred (self-censored) were it not for the operation of the platform*.

Too often, debates over intermediary liability completely elide this calculus. Online intermediaries almost certainly influence the deterrent effect of background legal rules on first-party actors, even if it is uncertain in which direction and to what magnitude that effect is felt. It is surely possible for online intermediaries to cooperate with law enforcement in ways that would magnify the law’s deterrent effect by increasing the likelihood that responsible actors will be held accountable. But the extent to which intermediaries would find it beneficial to maximize that cooperation likely hinges on what incentives and disincentives they face.<sup>103</sup>

Most notably, the relative anonymity (or pseudonymity) of online interactions mediated by platforms can serve to make direct enforcement of the law difficult or even impossible. Consider the case of short-term rental laws. As economists Jian Jia and Liad Wagman discuss, “[e]nforcement of [short-term rental laws] is made difficult because of a certain degree of anonymity that is afforded to

---

<sup>102</sup> Due to a combination of transaction costs for plaintiffs in continually monitoring and prosecuting cases against dispersed defendants and the likelihood that many users who generate illegal content are judgement-proof.

<sup>103</sup> Often it appears that online platforms have been quite willing to work with law enforcement by sharing user data with them. See Maggie Gile, *Big Tech Complied With 85% of Government Requests, Handed Over Data in First Half of 2020*, NEWSWEEK (Jun. 22, 2021), <https://www.newsweek.com/big-tech-complied-85-government-requests-handed-over-data-first-half-2020-1603070>. But there are also a few notable examples where such cooperation was not forthcoming, even to enable a court (as opposed to an enforcement agency) to properly exercise its jurisdiction under the All Writs Act, 28 U.S.C. § 1651. See, e.g., David W. Opperbeck & Justin (Gus) Hurwitz, *Apple v. FBI: Brief in Support of Neither Party in San Bernardino iPhone Case* (Mar. 10, 2016), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2746100](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2746100). In other cases online platforms have strenuously resisted enforcement agency efforts to obtain data allegedly relevant to harmful online activity. See, e.g., Ashby Jones, *The Government and Google*, WALL ST. J. (Jan. 19, 2006), <https://www.wsj.com/articles/BL-LB-239>; Geoffrey A. Manne, *If Government Is the Problem, When Is Google the Solution?*, TRUTH ON THE MARKET (Jan. 20, 2006), <https://truthonthemarket.com/2006/01/20/if-government-is-the-problem-when-is-google-the-solution/>.

hosts as part of what the platform often cites as protecting its users' privacy."<sup>104</sup> Similarly, in the *Oberdorf* case, Amazon's third-party merchant system made it marginally easier for a third party to remain anonymous and ultimately escape liability for a defective product.<sup>105</sup> While it is not impossible for the owner of a brick-and-mortar retail outlet to conduct business while remaining similarly anonymous, it is surely more difficult.<sup>106</sup> As Jia and Wagman go on to note, Section 230 has directly reinforced this dynamic and impeded efforts to enforce the law: "[T]he Communications Decency Act has been used by platforms to fortify the privacy—and thereby anonymity—of sellers, which makes enforcement of past and new regulations difficult."<sup>107</sup>

So what are the costs of this current legal regime, and what effects does it have on how online intermediaries behave? To be clear, the current legal regime does not offer online intermediaries absolute immunity from the threat of litigation or liability arising from hosted content.<sup>108</sup> According to the most comprehensive study of Section 230 cases, published by the Internet Association:

[F]ar from acting as a "blanket immunity" ... Section 230 only served as the primary basis for a ruling in 42 percent of the decisions we reviewed. When courts are concerned platforms may have played a role in creating content, they require discovery before deciding whether to grant 230 immunity. Our review also revealed that in many decisions, the underlying claims where defendants asserted a Section 230 defense were dismissed for lacking merit.<sup>109</sup>

---

<sup>104</sup> Jian Jia & Liad Wagman, *Platform, Anonymity, and Illegal Actors: Evidence of Whac-a-Mole Enforcement from Airbnb*, 63 J.L. & ECON. 729, 730 (2020). Importantly, this is no defense of those particular laws in the first place, about which one could make a strong case they are special interest protections for hotels and motels that harm consumer welfare. But the answer to bad laws is to repeal them, not to create an indiscriminate loophole on the Internet that happens to make them, as well as good laws, harder to enforce.

<sup>105</sup> See *Oberdorf v. Amazon.com, Inc.*, 930 F.3d 136, 142 (3rd Cir. 2019) ("Neither Amazon or Oberdorf has been able to locate a representative of the Furry Gang, which has not had an active account on Amazon.com since May 2016.").

<sup>106</sup> A key point here is that this concern in no way implicates anonymous speech in the sense that public transparency is required. Although that could be helpful in cases where the optimal method of preventing harmful conduct is self-help, for the most part this is a concern with ensuring that intermediaries have the ability to identify their third-party merchants when and if a court or law enforcement official requires it.

<sup>107</sup> Jia & Wagman, *supra* note 104, at 730. ("Measures have been proposed to hold platforms accountable for illegal listings, but these measures have faced strong and thus far successful legal resistance that cites the protections that platforms are afforded under the Communications Decency Act of 1996 and secured a recent Supreme Court victory in a related privacy battle over guests' information.") (referencing *City of L.A. v. Patel*, 135 S.Ct. 2443 (2015)).

<sup>108</sup> See generally Ardia, *supra* note 61. For a recent example of a court conducting a thorough analysis of a plaintiff's claims in order to determine whether the platform was actually the creator of the content at issue due to the tools it created (and thus not a beneficiary of Section 230 immunity), see *Vargas v. Facebook*, 19-cv-05081-WHO (N.D. Cal. Aug. 20, 2021).

<sup>109</sup> ELIZABETH BANKER, INTERNET ASSOCIATION, A REVIEW OF SECTION 230'S MEANING & APPLICATION BASED ON MORE THAN 500 CASES (July 27, 2020) at 2, available at <https://internetassociation.org/publications/a-review-of-section-230s-meaning-application-based-on-more-than-500-cases/> [hereinafter, "IA Section 230 Case Review"]. See also Neil Fried, *IA Study Shows Sec 230 Reform Would Have Impact Only Where Needed*, DIGITALFRONTIERS ADVOCACY: BLOGS & OP EDS (Aug. 3, 2020), <https://digitalfrontiersadvocacy.com/blogs-and-op-eds/f/ia-study-shows-sec-230-reform-would-have-impact-only-where-needed>.

As a result, “most courts conducted a careful analysis of the allegations in the complaint, and/or of the facts developed through discovery, to determine whether or not Section 230 should apply.”<sup>110</sup> For a sizeable number of cases, Section 230 did not cut litigation off at the earliest stage.<sup>111</sup> For these defendants, in other words, the absence of Section 230 might make little difference to the litigation they face. Of course, we do not know how many more cases would be brought without Section 230 (it is fair to surmise there would be more), nor whether the same pattern of outcomes would be observed. But we do know that some 70 percent of cases implicating Section 230 today would likely entail the same or only marginally higher litigation costs in the absence of Section 230.<sup>112</sup>

That such a high percentage of cases were not dismissed at the earliest stage, and that only 42 percent ultimately were decided on the basis of Section 230 immunity, means there is already incentive to bring platform-liability cases. In other words, it can hardly be argued that Section 230 fully deters potential plaintiffs from bringing suit against online intermediaries because they have no chance to prevail or no ability to impose costs on defendants sufficient to lead to a settlement. Obviously, Section 230 diminishes incentives to sue, but those incentives are manifestly not zero. Indeed, as another systematic review of Section 230 cases noted at the time of its publication in 2010:

While section 230 has largely protected intermediaries from liability for third-party speech, it has not been the free pass many of its proponents claim and its critics lament it to be. First, intermediaries continue to face legal claims arising from the speech of third parties. Indeed, the data show that plaintiffs have filed an increasing number of such cases each year. Second,

---

<sup>110</sup> IA Section 230 Case Review, *supra* note 109, at 7 (“Of the court decisions reviewed, courts relied primarily on the Section 230 immunity to determine the outcome in 42 percent of the decisions... even when courts applied Section 230 immunity at the motion to dismiss stage there are innumerable examples where courts gave plaintiffs multiple opportunities to amend complaints to try to avoid the Section 230 immunity.... [i]n our sample, courts refused to apply Section 230 immunity in over 12 percent of the decisions because there was an exception applied or the court determined that the immunity was not applicable to the case before it.”).

<sup>111</sup> *Id.*

<sup>112</sup> At least one scholar attributes this litigation reality to the absence of a fee-shifting mechanism in Section 230 and the lack of an efficient mechanism for subjects of online defamation to rebut untruthful claims. See Anthony Ciolli, *Chilling Effects: The Communications Decency Act and the Online Marketplace of Ideas*, 63 U. MIAMI L. REV. 137, 138 (2008) (“Congress’s failure to foresee an imminent change in the nature of the Internet... resulted in § 230 of the CDA failing to contain a provision authorizing the recovery of attorneys’ fees and court costs in litigation where § 230 was successfully used as a defense. Furthermore, Congress’s inability to predict the hegemony of Google’s algorithm-based search engine and the increasingly popular practice of “googling” potential employees, friends, and dates have hindered the development of an efficient mechanism for individuals to rebut untruthful information about themselves that has been preserved in perpetuity in Google and other search engines, and thus further encouraged the filing of frivolous lawsuits against immunized Internet intermediaries as a method of clearing one’s name.”). There is little evidence to support this claim, however, and the extensive literature on the consequences of different fee-shifting rules is studiously ambivalent. See Avery Wiener Katz & Chris William Sanchirico, *Fee Shifting*, in *PROCEDURAL LAW AND ECONOMICS, ENCYCLOPEDIA OF LAW AND ECONOMICS* 271 (Chris William Sanchirico ed., 2d ed. 2012) (“It is unclear whether fee shifting increases the likelihood of settlement, whether it decreases total expenditures on litigation or total payouts by defendants, or whether it on balance improves incentives for primary behavior. It is even unclear whether fee shifting makes it easier for parties with small meritorious claims to obtain compensation, in light of the increased costs per case that it induces.”).



even in cases where the court dismissed the claims, intermediaries bore their own legal costs, and it took courts nearly a year, on average, to issue a decision addressing the intermediary's defense under section 230. *Although section 230 set a high bar for plaintiffs to overcome, more than a third of their claims survived preemption.*<sup>113</sup>

In addition, where cases were dismissed early, many were dismissed for reasons other than Section 230 immunity.<sup>114</sup> For these cases—where defendants prevailed on such grounds as the First Amendment, anti-SLAPP statutes,<sup>115</sup> or because the plaintiff failed to state a cognizable claim—Section 230 is not doing any obvious work at all. While we do not know whether this same pattern would hold following a hypothetical removal or reform of Section 230, the law's presence does not appear to be decisive in determining the cost of litigation or the finding of liability.<sup>116</sup> It is noteworthy that, even with an apparently substantial incentive to sue and with literally billions of potential plaintiffs and an unfathomable volume of potentially tortious content, only about 500 cases related to Section 230 have been filed in the quarter-century since 1996.<sup>117</sup> This could be evidence that the law stands as a strong disincentive to bring suit, but it is at least plausible that fears of an avalanche of vexatious litigation in the absence of Section 230 immunity are overblown.

The same can also be said for the potential costs of moderation to avoid liability risks. Platforms today expend enormous resources on content moderation. Of course, they do so not out of legal obligation, but out of a belief that content moderation improves the quality of their services. Regardless of the motivation, however, there is surely some overlap between voluntarily moderated content and content likely to raise litigation risks. In the wake of Section 230 reform, how much of

---

<sup>113</sup> *Ardia, supra* note 61, at 382 (emphasis added).

<sup>114</sup> IA Section 230 Case Review, *supra* note 109, at 7 (“In over a quarter of decisions (28 percent), the courts dismissed claims without relying on Section 230 because the claims lacked merit or were flawed for another reason. More than 140 of the 516 decisions examined resulted in claims being dismissed in whole or in part by judges for failing to adequately plead legal violations without relying on Section 230.”). *See also* *Ardia, supra* note 61, at 493 (“In the majority of those decisions, however, the courts did not need to reach the question of section 230’s application because they found that the claims against the intermediary warranted dismissal on other grounds. When these decisions are included in the calculations, defendants won dismissal on section 230 or other grounds in more than three-quarters of the cases studied.”).

<sup>115</sup> *See e.g.,* Media Law Resource Center, *Anti-SLAPP Statutes and Commentary*, <https://www.medialaw.org/component/k2/item/3494> (explaining anti-SLAPP laws) (last visited June 1, 2021).

<sup>116</sup> *See, e.g.,* Brian L. Frye, *The Possible Redundancy of §230*, THE RECORDER (Nov. 12, 2017), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3069794](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3069794) (“Both fans and foes of *Zeran* assume that its interpretation of Section 230 changed the scope of liability for ISPs under the common law republication rule. I’m not so sure. While Section 230 requires courts to use different words than the common law rule, the *Zeran* interpretation of Section 230 produces essentially the same results as the common law rule, properly applied.”) The same is true of the limited number of cases implicating criminal activity. *See* IA Section 230 Case Review, *supra* note 109, at 10 (“Even where Section 230 was the basis for dismissal, the underlying facts and claims in the cases were largely identical to those dismissed on other grounds and just as easily could have been resolved based on the failure to state a claim as they were in the other cases.”).

<sup>117</sup> This may be inaccurate. It is difficult to tell from the IA Section 230 Case Review’s discussion of methodology whether the reviewed sample represents *all or virtually all* the cases since 1996 implicating Section 230, or simply a 516-case subset of a much larger number of cases. It does appear from the methodological discussion, however, that, at the very least, this likely represents the vast majority of cases. *See id.*, at 12.

the current expenditure would be diverted to different kinds of moderation? In other words, the relevant consideration is not the expected cost of litigation, in isolation; it is the expected cost of litigation *net of current expenditures that would serve to reduce the risk of litigation*.

In theory, [“conscript[ing] existing monitoring duties wholesale in the service of a new gatekeeper regime”] yields a novel duty with a well-understood and predefined focus. The fiction is that gatekeepers need not monitor more than they otherwise would; they need only monitor with an additional interest at stake (preventing the targeted misconduct) and an additional reason to perform carefully (expanded liability for breach). The fact, of course, is that additional liability generally increases monitoring costs and legal risks. Nevertheless, where a monitoring duty already exists, there may be attractive economies of scale in extending the range of its beneficiaries. Moreover, precisely because duties minted in this fashion are easily stated and prefocused, they are tempting vehicles for judicial innovation in tort, where they are often described as “relaxing” privity requirements that previously limited due care obligations to direct contractual relationships.<sup>118</sup>

It’s fair to assume that current moderation practices are not perfectly optimized to reduce litigation risk, but at least some (and perhaps most) of these practices would still be employed in a post-Section 230 world. And given that online intermediaries currently face liability for various causes of action that are exempt from Section 230, some existing moderation activities are already directed precisely toward mitigating liability risk.<sup>119</sup> If companies incur those costs even with the Section 230 shield in place, continuing to incur them in the absence of the law is not a cost properly attributed to reform.

Finally, these marginal costs would likely change over time in the years following Section 230 reform. The law is not static. It may be the case that, to a first approximation, a platform will face more litigation and liability costs the more content it hosts and/or the more moderation it performs. But as certain best practices become enshrined and recognized by the law, adherence to these could become effective safe harbors from liability, and even from litigation costs, particularly if fee-shifting outcomes track such practices. Whatever the “background” litigation/liability risk might be at the moment that immunity is relaxed, with increased legal certainty, it should fall over time.

## **B. Assumptions regarding the marginal change in litigation costs**

As mentioned above,<sup>120</sup> opponents of Section 230 reform have likely substantially overstated the marginal change in litigation costs that amending the law would engender. It is a truism that the more potential plaintiffs there are, the greater the likelihood of lawsuits, both meritorious and vexatious. Limiting immunity wouldn’t mean that every user will sue, but it surely means that more will. But the law and economics of litigation and civil procedure is much more nuanced than that.

---

<sup>118</sup> Kraakman, *Gatekeepers*, *supra* note 6, at 80.

<sup>119</sup> See 47 U.S. Code § 230(e).

<sup>120</sup> See *supra* Part II.B.

For one thing, there have been enormous changes to civil procedure since Section 230 was enacted in 1996. Most significantly, the Supreme Court decided *Bell Atlantic v. Twombly*<sup>121</sup> in 2007 and *Ashcroft v. Iqbal*<sup>122</sup> in 2009. Collectively, *Twombly/Iqbal* greatly reduced plaintiffs' ability to survive a motion to dismiss and get to the discovery stage of trial. Indeed, mitigating the burden of discovery on defendants was a crucial issue in the Court's decision in both cases:

The underlying issue animating... the Supreme Court's decisions in *Twombly* and *Iqbal*... is discovery access. When discovery costs are asymmetrically high for defendants, a plaintiff's ability to get through the answer/[motion to dismiss] stage can be a powerful club. Liberal pleading rules may have an *in terrorem* effect on defendants in these cases, possibly inducing more, and more one-sided, settlements. The Supreme Court's opinions in both *Twombly* and *Iqbal* take note of this point; the opinions make repeated and extensive references to the burden of discovery borne by large corporations (as in *Twombly*) or government officials (as in *Iqbal*).<sup>123</sup>

Although the precise magnitude of the “negative effect”<sup>124</sup> on plaintiffs of *Twombly/Iqbal* is difficult to measure, the most rigorous assessment to date, by economist Jonah Gelbach, estimates a lower bound on the effect. For civil rights cases, “*Twombly/Iqbal* negatively affected plaintiffs in at least... 18.1% of cases...” and, for cases not involving civil rights, “*Twombly/Iqbal* negatively affected at least 21.5% of plaintiffs facing [motions to dismiss.]”<sup>125</sup> These are substantial effects that suggest the duck-bites problem has been at least somewhat ameliorated since Section 230 was enacted.

There are good reasons to suspect that our complex morass of civil-procedure rules—significantly influenced by the plaintiffs' bar—still results in excessive litigation, as well as excessive expenditure on discovery and other aspects of the adversarial trial system.<sup>126</sup> But looking solely at the number of cases or even the total costs of litigation does not reveal whether either is “excessive.”

---

<sup>121</sup> 550 U.S. 544 (2007) (overturning the liberal notice pleading regime that had governed the motion-to-dismiss standard since 1957. See *Conley v. Gibson*, 355 U.S. 41 (1957)).

<sup>122</sup> 556 U.S. 662 (2009).

<sup>123</sup> Jonah B. Gelbach, *Locking the Doors to Discovery? Assessing the Effects of Twombly and Iqbal on Access to Discovery*, 121 YALE L.J. 2270, 2286 (2012).

<sup>124</sup> See *id.* at 2278 (“Taken together, discovery-prevented cases and settlement-prevented cases constitute the set of what I call ‘negatively affected cases,’ because these are cases whose disposition leads to worse results for the plaintiffs who bring suit.”).

<sup>125</sup> *Id.* at 2278-79. For reasons of proper study design, this latter group technically excludes not only civil rights cases, but also employment discrimination cases and cases involving financial instruments.

<sup>126</sup> Among the most well-known of the many academic books and articles making this claim, see, e.g., WALTER K. OLSON, *THE RULE OF LAWYERS: HOW THE NEW LITIGATION ELITE THREATENS AMERICA'S RULE OF LAW* (2002); RICHARD A. EPSTEIN, *SIMPLE RULES FOR A COMPLEX WORLD* (1995); George L. Priest, *The Modern Expansion of Tort Liability: Its Sources, Its Effects, and Its Reform*, 5 J. ECON. PERSP. 31 (1991); Kevin M. Murphy, Andrei Shleifer & Robert W. Vishny, *The Allocation of Talent: Implications for Growth*, 106 Q.J. ECON. 503 (1991); WALTER K. OLSON, *THE LITIGATION EXPLOSION: WHAT HAPPENED WHEN AMERICA UNLEASHED THE LAWSUIT* (1991); PETER W. HUBER, *LIABILITY: THE LEGAL REVOLUTION AND ITS CONSEQUENCES* (1988); Derek C. Bok, *A Flawed System of Law Practice and Training*, 33 J. LEGAL EDUC. 570 (1983).

It is well-understood that there are “fundamental differences between private and social incentives to use the legal system. These differences permeate the litigation process, from the choice of a harmed party whether to bring suit, to the plaintiff’s and the defendant’s negotiation over settlement versus trial, to their various decisions about legal expenditures.”<sup>127</sup> What remains unknown is the magnitude of divergence between social and private incentives, or even the direction of the net effect. “As a result, the privately determined level of litigation can be either socially excessive or socially inadequate.”<sup>128</sup>

Potential plaintiffs often will not take account of all relevant social costs when choosing both whether to bring suit and what type of suit to bring. This can lead to excessive and overly expensive litigation.<sup>129</sup> But there are also *benefits* of litigation that potential plaintiffs don’t internalize, potentially leading to too few suits and inefficiently small litigation expenditures:

But there is another important source of divergence between the private and the social incentive to use the legal system: that involving the difference between the private and the social *benefits* of its use. This divergence in benefits can work either to exacerbate or to counter the tendency toward its excessive use due to the private-social cost divergence. To explain, consider one of the principal social purposes of litigation, deterrence of unwanted behavior. This social goal has little to do with a person’s decision whether to bring suit. The motive of a person who brings suit is ordinarily not chiefly, if at all, to deter socially undesirable behavior in the future. Rather it is usually to obtain compensation for harm or other relief. Therefore, the plaintiff’s benefit from suit does not bear a close connection to the social benefit associated with it and may bear almost no connection at all.... If the private benefit falls short of the social benefit, however, there may be too little incentive to bring suit.<sup>130</sup>

Crucially, this assessment relies on the implicit understanding that litigation has positive social functions—not the least of which is the deterrence of tortious behavior (and thus, ultimately, further lawsuits). The fact that defendants must bear the costs of litigation is a feature of the system, not a bug. Indeed, getting the legal incentives right requires that defendants pay for the costs of

---

<sup>127</sup> Steven Shavell, *The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System*, 26 J. LEG. STUD. 575, 577 (1997).

<sup>128</sup> Jonah B. Gelbach & Bruce H. Kobayashi, *The Law and Economics of Proportionality in Discovery*, 50 GA. L. REV. 1093, 1100 (2016). *See also id.* at 1102 (“Since external social costs are associated with too much litigation, while external social benefits are associated with too little, there are gross effects operating in both directions. As a matter of simple arithmetic, then, the net impact of these gross effects might point in either direction. Thus, whether there is too much, too little, or just the right amount of litigation in general is not a conclusion that can be drawn on a priori grounds.”).

<sup>129</sup> *See* Shavell, *Fundamental Divergence*, *supra* note 127, at 578 (“Specifically, when a person brings suit, he bears only his own legal expenses; he does not take into account that his suit will cause the defendant and possibly the court to incur legal expenses as well; a bias toward excessive suit is thus engendered. Similarly, once suit has been brought, when either litigant considers making a particular expenditure on litigation, he will not count as a cost to himself the expense that the opposing side and the court may be forced to bear as a consequence; this leads to an excessive level of litigation expenditures.”).

<sup>130</sup> *Id.*

litigation: “[D]efendants should pay for more than just the harm that they cause in adverse events: they should incur a bill equal to harm plus total litigation costs.”<sup>131</sup>

Once the benefits of litigation are allowed into the calculus, it becomes clear that discovery, while costly, likewise plays an important role in social welfare:

First, since discovery requests impose costs borne by the responder, some of our discovery system’s costs are externalized. In some cases, this effect may predominate. When and where it does, limitations on discovery-as-of-right, such as the proportionality standard, might be worth imposing. Second, though, *it is important to remember that in some cases, discovery will create social benefits by inducing revelation of evidence that yields socially beneficial litigation outcomes.*<sup>132</sup>

As noted, defenders of the status quo often assume there would be *no* benefits from litigation against platforms. Critiques of “excess” litigation costs are thus an extreme version of the claim that litigation is inefficient if it costs more than the amount of harm at issue in the suit (which may sometimes be the case, as many litigation costs, including discovery, may prove relatively fixed and may not reflect the magnitude of the harm alleged). But “it is perfectly possible for more to be spent on suit than the amount in question and yet for suit to be socially desirable. Indeed, it may be desirable for the state to encourage litigation even when total litigation costs exceed the amount at issue.”<sup>133</sup> The reason is that any given lawsuit may engender substantial deterrence (and, thus, outsized social benefit). While the costs of litigation are borne only when harm occurs, the costs of precaution may be much smaller. If litigation today leads to proper precautions—and thus, the absence of both future harms and future litigation—the net social savings may be substantial, even where the present litigation proves far costlier than the magnitude of harm at issue in the case.<sup>134</sup>

To the extent that the civil-law process yields society’s intended results, the threat of litigation may induce proper precautions and not result in excessive social costs. Indeed, under a negligence rule, “since non-negligent behavior often means that suit will not be brought (or that a claim will quickly be dropped)... the problem of socially excessive suit may thus be of less significance under the negligence rule than under strict liability.”<sup>135</sup> It is true that absolute immunity would lower

---

<sup>131</sup> *Id.* at 579 (“The reason is that, when an injurer causes harm and is sued, the true cost that society is forced to bear equals the harm plus total litigation costs. Accordingly, for injurers’ incentives to reduce risk to be appropriately strong, they need to pay for more than only the direct harm that they cause.”).

<sup>132</sup> Gelbach & Kobayashi, *supra* note 128, at 1103 (emphasis added).

<sup>133</sup> Shavell, *Fundamental Divergence*, *supra* note 127, at 584.

<sup>134</sup> *Id.* (“[N]o litigation costs will actually be borne, yet injurers will be led to act appropriately. Specifically, if injurers know that they will definitely be sued for negligently causing harm, they will be induced to act non-negligently. And since our supposition is that courts will never make a mistake in assessing negligence, there would never be any findings of negligence, so that victims would never bring suit.”).

<sup>135</sup> *Id.* at 598-99.

litigation costs further still, but at the expense of the social benefits of deterrence. A few expensive lawsuits in exchange for optimal deterrence could easily be a net social benefit.

We don't know, of course, how large this effect might be. But it certainly exists, and almost certainly is quite large. We know this because platforms already spend considerable resources on moderation, even with significantly less risk of liability or litigation costs. We don't know that these moderation costs line up perfectly with what would be incurred under a post-Section 230 liability regime, but certainly there is substantial overlap.<sup>136</sup> Indeed, there is nearly perfect overlap when it comes to federal criminal law, intellectual property, and sex trafficking, because these areas are explicitly exempted from Section 230 immunity, and platforms already face the risk of lawsuits over these issues. The bottom line is that, whatever total liability costs platforms would bear under a more limited intermediary-immunity regime, some portion of those costs are already born by platforms today.

It is also essential to understand that precautions like content moderation are a substitute for litigation. "To induce injurers to exercise the proper level of precautions, a level that reflects the full measure of social costs that are incurred when they cause harm and are sued, the amount that they pay must equal the direct harm that they cause plus the sum of all litigation costs."<sup>137</sup> But this is true even where defendants don't directly cause harm, but could nevertheless cost-effectively avoid it.

What determines this is the extent to which the threat of lawsuit induces potential defendants to take precautions that deter tortious conduct. Because platforms already undertake considerable moderation, it may be that there is little additional deterrence to be gained from increasing liability risk and litigation costs. It's also possible that too much potentially tortious conduct isn't recognizable or identifiable *ex ante*, such that there is nothing practical that platforms could do to avoid it. In these cases, it's possible that the gains from additional liability/litigation may be negligible.

But they also may not be. Online platforms realize benefits from tortious content even when it harms others, thus undermining their incentive to police content at the socially optimal level. Further, the transaction costs of pursuing case-by-case litigation against users may be so large that platforms are currently under-incentivized to protect against these harms. If the costs were better internalized by platforms, it's possible that they would be willing and able to efficiently prevent them. It is in precisely such circumstances that the law and economics literature supports indirect liability—

---

<sup>136</sup> Thus, it is also worth noting that in a post-reform world it could be the case that platforms incur additional compliance costs over and above their baseline moderation costs. But, again, this is not necessarily a negative if the compliance costs help to better align platform behavior with the socially optimal outcome.

<sup>137</sup> Shavell, *Fundamental Divergence*, *supra* note 127, at 588.

“where liability would serve to encourage a party to internalize some significant negative externality unavoidably associated with its activities.”<sup>138</sup>

Thus, platforms may know that the harms their users cause are usually too minor for a victim to have incentive to pursue redress, even if the aggregate cost of such harms exceeds the cost the platform would incur to take action to avoid them. It’s also plausible that the difficulties users face in meeting the requisite legal standards to bring suit are so large that they could not bring action in response even to harms of large magnitude, which also would not be deterred by platforms. And, of course, certain harms—such as civil rights abuses—may entail a combination of low monetary awards and difficulty proving liability such that these suits are rarely brought, even though the effect of prevailing might be socially significant.

### C. Common law antecedents to an intermediary duty of care

The common law has long embraced the notion of indirect or vicarious liability, precisely for the purpose of aligning incentives where they can be most useful:

[R]ules that hold one party liable for wrongs committed by another are the standard legal response in situations where... liability will be predictably ineffective if directly applied to a class of bad actors and yet there exists a class of related parties capable of either controlling those bad actors or mitigating the damage they cause.... [W]hile indirect liability comes in a wide variety of flavors and forms..., it is the norm.<sup>139</sup>

As Lichtman & Posner note in the article quoted above, the economic analysis undergirding this sort of common-law liability quite obviously applies online: “Our argument in favor of... liability is primarily based on the notion that [online intermediaries] are in a good position to reduce the number and severity of bad acts online.”<sup>140</sup>

It is impossible to know exactly how a robust common law of online intermediary liability would have developed in a world where Section 230 immunity never existed. But it is informative to consider how the offline world has dealt with third-party liability, especially when an intermediary operates under a duty of care with respect to third parties. Supporters of the status quo have argued that the common law duty of reasonableness would be a poor fit for online intermediaries.<sup>141</sup> But

---

<sup>138</sup> Lichtman & Posner, *supra* note 24, at 230. Notably, Lichtman & Posner advocate for intermediary liability for ISPs, arguing that they “should to some degree be held accountable when their subscribers originate malicious Internet code, and... when their subscribers propagate malicious code. *Id.* at 222.

<sup>139</sup> *Id.* at 223.

<sup>140</sup> *Id.* On the law and economics of third-party liability, see, for example, Kraakman, *Gatekeepers*, *supra* note 6; Alan Sykes, *An Efficiency Analysis Of Vicarious Liability Under the Law of Agency*, 91 YALE L.J. 168 (1981); Alan Sykes, *The Economics of Vicarious Liability*, 93 YALE L.J. 1231 (1984); Reinier Kraakman, *Third Party Liability*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 673 (Peter Newman, ed. 1998).

<sup>141</sup> See, e.g., Enrique Armijo, *Reasonableness as Censorship: Section 230 Reform, Content Moderation, and the First Amendment*, 73 FLA. L. REV. (forthcoming 2021), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3764061](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3764061).

even though the contours of offline intermediary liability offer an inexact model, it is nevertheless instructive to explore how the law works in this space.

This Part will introduce principles from the common law for two main reasons: first, to illustrate how Section 230 immunity is a departure from normal rules governing intermediary behavior and, second, to provide a set of analogies for how these principles could operate online. Thus, although there is currently no real common law of online intermediary liability, there are antecedents and offline analogues from which lessons can be drawn regarding how a duty of care for online intermediaries could function.

Generally speaking, the law of negligence has evolved a number of theories of liability that apply to situations in which one party obtains a duty of care with respect to the actions of a third party.<sup>142</sup> One legal obligation of every business is to take reasonable steps to curb harm from the use of its goods and services. There are two bases for such a duty. If the business has created a situation or environment that puts people at risk, it has an obligation to mitigate the risk it has created. Secondly, if the business has entered into a relationship with someone, such as a potential customer it has invited onto its premises, it can have an affirmative obligation to prevent the risk of harm to that person even if the business did not directly create the risk.<sup>143</sup>

There are many examples at common law of the duty that business owners owe to their customers (or, sometimes, to the outside world) that analogize to online intermediaries. Owners of hotels, for instance, owe a reasonable duty of care to their paying guests when the owners are aware that a third party is victimizing or will victimize those guests. Georgia courts base this liability on “the proprietor’s superior knowledge of the perilous instrumentality and the danger therefrom to persons going upon the property.”<sup>144</sup> Further, “it is the duty of a proprietor to protect an invitee or guest from injury caused by a third person if the host is reasonably aware of the probability or likely possibility of such an act by a third party and such injury could be avoided or prevented by the host through the exercise of ordinary care and diligence.”<sup>145</sup>

---

<sup>142</sup> See, e.g., Restatement (Second) of Torts § 314A (1965) (“(1) A common carrier is under a duty to its passengers to take reasonable action (a) to protect them against unreasonable risk of physical harm, and (b) to give them first aid after it knows or has reason to know that they are ill or injured, and to care for them until they can be cared for by others. (2) An innkeeper is under a similar duty to his guests. (3) A possessor of land who holds it open to the public is under a similar duty to members of the public who enter in response to his invitation. (4) One who is required by law to take or who voluntarily takes the custody of another under circumstances such as to deprive the other of his normal opportunities for protection is under a similar duty to the other.”).

<sup>143</sup> See *Fowler V. Harper* and *Posey M. Kime*, *The Duty to Control the Conduct of Another*, 43 YALE L.J. 886, 887 (1934). See also DAN B. DOBBS ET AL., *HORNBOOK ON TORTS* § 9.2, at 191, § 20.1, at 459-60, § 20.6, at 465-66, § 25.1, at 615-16, § 25.4, at 620-21, §§ 26.1-26.5, at 633-44, §§ 26.9-26.10, at 651-55 (2d ed. 2015) (stating that a business has a duty to take reasonable steps to prevent one person from using its auspices to harm another if the business has a relationship with either party, such as by welcoming one or the other to engage with it, and that a failure to meet that duty can lead to liability).

<sup>144</sup> *Emory University v. Duncan*, 182 Ga. App. 326, 328, 355 S.E.2d 446 (1987).

<sup>145</sup> *Donaldson v. Olympic Health Spa*, 175 Ga. App. 258, 261, 333 S.E.2d 98 (1985).



Indeed, “innkeeper liability”—an obligation to ensure the safety and protection of guests—goes back centuries, and was premised on the assumption that the owner of a premises has control of the building, workers, and related facilities.<sup>146</sup> Further, the premises owner has some ability to select which guests he allows,<sup>147</sup> which further empowers him to afford protection to his guests from the bad acts of other guests.<sup>148</sup> Thus, an “innkeeper is bound to exercise reasonable care in protecting the guest within his inn from personal injury while remaining as his guest.”<sup>149</sup> Notably, the common law did not cast innkeepers as “insurers” of their guests, but merely obligated them to act reasonably in treating those guests.<sup>150</sup>

Over time, this doctrine has evolved differently in different courts, but remains essentially intact in its general outline. English law evolved similar duties for innkeepers, requiring them to take “reasonable care to prevent damage to [a] guest from unusual danger.”<sup>151</sup> This includes an obligation to reasonably protect guests from the foreseeable criminal acts of third parties.<sup>152</sup> Similarly, in the United States, the fact of third-party criminal acts does not relieve those with a duty of care from liability if it is within the scope of the risk created by the negligence.<sup>153</sup>

Examples of premises liability for businesses are similarly instructive, and the Restatement (Second) of Torts describes a similar theory for such liability:

A possessor of land who holds it open to the public for entry for his business purposes is subject to liability to members of the public while they are upon the land for such a purpose, for physical harm caused by the accidental, negligent, or intentionally harmful acts of third persons or animals, and by the failure of the possessor to exercise reasonable care to (a) discover that such acts are being done or are likely to be done, or (b) give warning adequate to enable the visitors to avoid the harm, or otherwise protect them against it.<sup>154</sup>

---

<sup>146</sup> Joseph James Hemphling, *Innkeeper’s Liability at Common Law and under the Statutes*, 4 NOTRE DAME L. REV. 421, 427 (1929).

<sup>147</sup> Notably, hotels and inns differ from intermediaries insofar as hotels have common carrier obligations—they can only *reasonably* refuse guests, not absolutely refuse them. Arguably, online intermediaries have a much stronger ability to refuse access to their services as they are not subject to common carrier obligations.

<sup>148</sup> Hemphling, *supra* note 146, at 427-28

<sup>149</sup> *Id.* at 428.

<sup>150</sup> *Id.*

<sup>151</sup> *Al-Najar & Ors v. The Cumberland Hotel LTD*, EWHC 1593, Queen’s Bench Division, at para. 184 (2019), *available at* <https://www.casemine.com/judgement/uk/5d10902b2c94e063b32426f8>.

<sup>152</sup> *Id.* at ¶ 195.

<sup>153</sup> See Restatement (Second) of Torts § 442(b) (1965) (“Where the negligent conduct of the actor creates or increases the risk of a particular harm and is a substantial factor in causing that harm, the fact that the harm is brought about through the intervention of another force does not relieve the actor of liability, except where the harm is intentionally caused by a third person and is not within the scope of the risk created by the actor’s conduct.”).

<sup>154</sup> Restatement (Second) of Torts § 344 (1965).

Shopping malls owe a duty of care to invitees who come to shop at their various stores.<sup>155</sup> They must ensure invitees are protected from foreseeable harms while within the mall, including those committed by third parties.<sup>156</sup> This means, among other things, providing trained security personnel, well-lit parking lots, and mechanisms to make sure that stores within the mall do not defraud customers or sell them defective products.<sup>157</sup>

Private parks give rise to similar duties of care that provide a useful model to consider. Much like public parks, private parks are areas where people can gather and hold social events, camp, or enjoy nature and other outdoor activities. The operators of private parks owe a duty of care to individuals who come onto the property.<sup>158</sup> Park owners have a duty to invitees who paid or were invited to enter to ensure their safety during their stay.<sup>159</sup> This includes keeping the property reasonably free of dangerous conditions by performing regular inspections and fixing hazards in a reasonable amount of time, or posting clearly visible warnings.<sup>160</sup> If reasonable diligence would discover a hazard, then property owners can be held liable for failing to mitigate it or warn about it.<sup>161</sup>

For licensees, which would include guests who come onto the property even without any potential commercial relationship, owners must exercise reasonable care to prevent foreseeable injuries.<sup>162</sup> There are instances where a property owner allows use without any consideration (i.e., without payment).<sup>163</sup> In that context, they still must notify licensees about known hazards, but do not have a specific duty to inspect premises before allowing entry.<sup>164</sup>

The common law has also recognized situations where there is a duty to control the conduct of certain actors to prevent them from causing harm to another when “a special relation exists

---

<sup>155</sup> See, e.g., Krista Sheets, *Mall Security Duties and Law: Shopping Mall Injury Attorney*, LEGALMATCH (Aug. 8, 2020), <https://www.legalmatch.com/law-library/article/shopping-mall-liability-lawyers.html>.

<sup>156</sup> See *id.*

<sup>157</sup> See *id.*

<sup>158</sup> See, e.g., *Here's Who Can Be Held Responsible For Your Injuries In A Park (Private vs. Public)*, COMPASS LAW GROUP, LLP (Jun. 21, 2018), <https://cmplawgroup.com/blog/heres-who-can-be-held-responsible-for-your-injuries-in-a-park-private-vs-public/>.

<sup>159</sup> See *id.*

<sup>160</sup> See Michael Waks, *Invitees, Licensees & Trespassers: What's the Difference?*, LAW OFFICE OF MICHAEL D. WAKS (Jun. 14, 2019), <https://www.michaelwaks.com/invitees-licensees-trespassers/>.

<sup>161</sup> See *id.*

<sup>162</sup> See *id.*

<sup>163</sup> See *id.*

<sup>164</sup> See *id.*

between the actor and the third person.”<sup>165</sup> So, in some instances, the common law imposes a duty to restrain another person’s conduct in order to protect third parties or the public.<sup>166</sup>

The traditional duty to control another arose most commonly in cases of jailors controlling prisoners or mental health professionals protecting the public from particularized threats from patients under their care.<sup>167</sup> But courts have also imposed duties in other situations with special relationships. For instance, one Tennessee court found a duty for an adult hosting a party of underage teens who were drinking alcohol:

The duty of care [the defendant] owed to his guests, however, lies separate and apart from furnishing alcohol. Because he knowingly permitted and facilitated the consumption of alcohol by minors, an illegal act, [the defendant] had a duty to exercise reasonable care to prevent his guests from harming third persons.<sup>168</sup>

The court analogized the situation to previous cases where doctors and hospitals were found to have a duty to warn about communicable diseases or the effects of medication when there was foreseeable risk of harm to others.<sup>169</sup>

Dram shop liability presents a particularly apt analogy. Dram shop liability refers to the legal duty of care that applies to the owners of taverns, bars, and similar establishments to protect others from the tortious acts of inebriated patrons.<sup>170</sup> U.S. courts were traditionally wary of imposing liability on tavern owners for the behavior of inebriated patrons who committed torts once they left the premises.<sup>171</sup> Some states’ common law did develop responsibilities owed to third parties or to

---

<sup>165</sup> Restatement (Second) of Torts § 315 (1965). This is a long-standing exception to the general rule that there is no responsibility to protect or control third persons. See *Fowler V. Harper & Posey M. Kime*, *The Duty to Control the Conduct of Another*, 43 YALE L.J. 886, 887 (1934) (noting that “certain socially recognized relations exist which constitute the basis for such legal duty”).

<sup>166</sup> Kenneth S. Abraham & Leslie Kendrick, *There’s No Such Thing as Affirmative Duty*, 104 IOWA L. REV. 1649, 1633 (2019) (“When the defendant has a special relationship that involves a duty to control the conduct of another, tort law sometimes imposes a duty to protect third parties against risks posed by the person controlled.”).

<sup>167</sup> *Id.* See also *Tarasoff v. Regents of Univ. of Cal.*, 551 P.2d 334, 340 (Cal. 1976) (on mental health professionals having a duty to control patients).

<sup>168</sup> *Biscan v. Brown*, 160 S.W.3d 462, 482 (Tenn. 2005).

<sup>169</sup> *Id.* at 479 (“We concluded a physician owes a duty to the immediate family members of a patient to warn them of possible exposure to the source of the patient’s illness, even in the absence of a physician-patient relationship with the immediate members of the family. Our holding rested on the fact that it was highly foreseeable that the patient’s wife would also contract the disease which killed the patient. Similarly... we held that a physician owed a duty to a third party as a member of the “motoring public” to warn his patient that medication could impair the patient’s driving ability because the patient’s medical history and the effects of medication made the third party’s injury foreseeable. In [another case,] this Court held that the defendant hospital owed a duty to the patient’s husband and to the general public to inform the patient that she had HIV because it was foreseeable that identifiable third parties would be at risk for exposure.”) (citations omitted).

<sup>170</sup> See 137 Am. Jur. Proof of Facts 3d 195 (Originally published in 2013).

<sup>171</sup> See, e.g., *State for Use of Joyce v. Hatfield*, 197 Md. 249, 254, 78 A.2d 754, 756 (1951); *Holmes v. Circo*, 196 Neb. 496, 500-01, 244 N.W.2d 65, 68 (1976); *Williamson v. Old Brogue, Inc.*, 232 Va. 350, 354, 350 S.E.2d 621, 624 (1986).

the intoxicated themselves, but the doctrine grew unevenly.<sup>172</sup> To date, at least 14 states have found some form of tort liability is appropriately assigned to the seller of intoxicating beverages when inebriated patrons subsequently commit tortious acts.<sup>173</sup>

Dram shop law has evolved largely as a consequence of state legislatures expressly shifting liability onto the owners of establishments based on their sale of intoxicating beverages.<sup>174</sup> The basic logic of dram shop liability, however, tracks in many respects the premises-liability case law. Tavern owners are least-cost avoiders; they are uniquely positioned to control how much alcohol they sell to a particular patron, and to cut that patron off when they appear to become intoxicated. The application is not always straightforward, with some courts construing dram shop laws as strict liability statutes, and others interpreting them through a negligence-style “reasonableness” lens.<sup>175</sup>

Thus, when a firm holds itself out to the public as providing a service, it does so “under a duty to exercise reasonable care toward those who benefit from the service.”<sup>176</sup> This duty can even extend to impose obligations relating to non-customers. As the Restatement (Second) of Torts puts it:

One who undertakes, gratuitously or for consideration, to render services to another which he should recognize as necessary for the protection of a third person or his things, is subject to liability to the third person for physical harm resulting from his failure to exercise reasonable care to [perform] his undertaking, if

- (a) his failure to exercise reasonable care increases the risk of such harm, or
- (b) he has undertaken to perform a duty owed by the other to the third person, or

---

<sup>172</sup> See Francis A. King, *Common Law Liability of the Liquor Vendor*, 18 CASE W. RESRV. L. REV. 251 (1966) (noting that at common law, when liability was found it was due to the particularly fragility or susceptibility of the intoxicated person, and that “able bodied men” becoming intoxicated would break the proximate causation that would enable suing a tavern owner). See also 97 A.L.R.3d 528 (Originally published in 1980) (“At common law, it was not a tort to either sell or give intoxicating liquor to ordinary able-bodied men, and no cause of action existed against one furnishing liquor in favor of those injured by the intoxication of the person.”).

<sup>173</sup> See 97 A.L.R.3d 528 (Originally published in 1980) (Noting cases over the course of the middle- and late-twentieth century that abrogated the prior common law rule against holding tavern owners liable in negligence for the tortious acts of intoxicated patrons).

<sup>174</sup> § 9:87. Liability under dram shop, or civil liability, acts, 2A American Law of Torts § 9:87. To date, thirty states have some form of dram shop liability statute in place. <https://www.ncsl.org/research/financial-services-and-commerce/dram-shop-liability-state-statutes.aspx>.

<sup>175</sup> *Id.* The difference is important practically as well as conceptually. Under a strict liability approach, the tavern owner is treated as something like an insuring party for the harms caused by patrons, whereas under a negligence approach, the tavern owners is held liable only when they failed to act reasonably.

<sup>176</sup> Abraham & Kendrick, *supra* note 166, at 1656.

(c) the harm is suffered because of reliance of the other or the third person upon the undertaking.<sup>177</sup>

Indeed, U.S. courts have increasingly allowed cases against online intermediaries to proceed undisturbed by Section 230 under state products-liability laws.<sup>178</sup> Section 230 has been found inapplicable in cases where the complaint is framed around the duty of care owed to users of a service. *Oberdorf v. Amazon*, to take one recent example of this trend, framed the case of a woman injured by a product purchased on Amazon.com as a products-liability tort.<sup>179</sup> In *Oberdorf*, a woman lost sight in one eye because Amazon facilitated a sale with a third party who sold a defective product. That seller then disappeared into the wind.<sup>180</sup> Amazon tried and failed to raise a Section 230 defense.<sup>181</sup> The Third Circuit construed the case as one about Amazon's *conduct* as opposed to being focused on the *speech* of the third party. Thus, by designing its system in such a way that users could not reliably contact third-party sellers, Amazon.com could be held liable under Pennsylvania products-liability law and unable to apply a Section 230 defense.<sup>182</sup>

The Ninth Circuit has been developing similar doctrine for several years. In *Model Mayhem* the court held that defendants had a duty to warn users about third parties who were known to use its services to locate victims and sexually assault them.<sup>183</sup> The court relied on California's duty-to-warn law to construe the relevant duty of care.<sup>184</sup> In that case, the plaintiffs alleged that defendants had actual knowledge that two of its users were using their service to locate victims.<sup>185</sup> Critically, Section 230 was unavailable as a defense because the website was not being sued for the speech acts of third parties, but for its own unreasonable failure to meet the duty of care it owed its users.<sup>186</sup>

Even more recently, in *Lemmon v. Snap*, the Ninth Circuit did not permit Section 230 to be raised as a bar to a defective-product-design case in which plaintiffs alleged that Snap had negligently designed its platform, Snapchat, in a way that encouraged the plaintiffs' children to drive at excessive

---

<sup>177</sup> Restatement (Second) of Torts § 324A (1965).

<sup>178</sup> See, e.g., *Loomis v. Amazon.com, LLC*, No. B297995 (Ct. App. 2nd Dist., Apr. 21, 2021), available at <https://www.bloomberglaw.com/public/desktop/document/KISHALOOMISPlaintiffandAppellantvBC632830AMAZONCOMLLCDefendantand?1621264076>.

<sup>179</sup> See *Oberdorf*, 930 F.3d at 136.

<sup>180</sup> See *id.* at 142.

<sup>181</sup> See *id.* at 151-53.

<sup>182</sup> See *id.* at 153-54.

<sup>183</sup> *Jane Doe 14 v. Internet Brands, Inc.*, 824 F.3d 846 (9th Cir. 2016).

<sup>184</sup> See *id.* at 850.

<sup>185</sup> See *id.* at 851 ("Instead, Jane Doe attempts to hold Internet Brands liable for failing to warn her about information it obtained from an outside source about how third parties targeted and lured victims through Model Mayhem. The duty to warn allegedly imposed by California law would not require Internet Brands to remove any user content or otherwise affect how it publishes or monitors such content.").

<sup>186</sup> See *id.* at 850-54.

speeds, which then led to their death in an automobile accident.<sup>187</sup> According to the court, “while providing content-neutral tools does not render an internet company a ‘creator or developer’ of the downstream content that its users produce with those tools, our case law has never suggested that internet companies enjoy absolute immunity from all claims related to their content-neutral tools.”<sup>188</sup> The court drew a clear distinction between claims based on the speech of third parties that would treat a platform as a publisher or distributor of user content, and claims based on product-design faults that treat the platform as a manufacturer:

Manufacturers have a specific duty to refrain from designing a product that poses an unreasonable risk of injury or harm to consumers... Meanwhile, entities acting solely as publishers—*i.e.*, those that “review[] material submitted for publication, perhaps edit[] it for style or technical fluency, and then decid[e] whether to publish it,”—generally have no similar duty.<sup>189</sup>

Thus, because the plaintiffs’ complaint did not seek to hold the defendant liable as a publisher or speaker, Section 230 could not be used to bar a negligent-design suit.<sup>190</sup>

One of the most relevant examples of an analogous standard of care arises in the context of defamation itself. Indeed, this analogy is discussed in *Barnes v. Yahoo*, one of the seminal Section 230 cases. “Defamation law sometimes imposes ‘an affirmative duty to remove a publication made by another.’”<sup>191</sup> In particular, a court may hold certain (offline) intermediaries liable when, “after knowledge of its existence, [an intermediary] negligently allow[s] the defamatory matter to remain for so long a time as to be chargeable with its republication.”<sup>192</sup> Although, unlike in our own proposal, an offline intermediary may be liable for defamation in such circumstances, the principle is similar. Indeed, in *Barnes v. Yahoo*, this doctrine was referenced as an analogy to the plaintiff’s claim in that case, which was framed not as a claim for defamation, but for “negligent provision or non-provision of services.”<sup>193</sup> Of course, as in that case—in which the court granted the defendant’s motion to dismiss based on Section 230 immunity—the distinction between a claim framed as a negligent failure to take care to prevent a tortious act and one framed as the commission of a tortious act itself may be somewhat blurry.

These and other examples of third-party liability under the common law provide useful guidance for understanding when such a cause of action might be appropriate in the online context:

---

<sup>187</sup> See *Lemmon v. Snap, Inc.*, Slip Op., No. 20-55295 (9th Cir. May 4, 2021).

<sup>188</sup> *Id.* at \*16.

<sup>189</sup> *Id.* at \*12 (internal citations omitted).

<sup>190</sup> See *id.* at \*17.

<sup>191</sup> *Barnes v. Yahoo!, Inc.*, 565 F.3d 560, 567 (9th Cir. 2009) (quoting Prosser and Keaton on Torts § 113, at 803).

<sup>192</sup> *Hellar v. Bianco*, 111 Cal.App.2d 424, 244 P.2d 757, 758 (1952) (cited in *Barnes v. Yahoo!*, *id.*)

<sup>193</sup> See Restatement (Second) of Torts § 324A (1965); see also *supra* note 177 and accompanying text.

Just as a delivery service might be held liable for delivering a package that obviously contains a ticking bomb, or a landlord might be held liable for permitting a use of his premises that is overtly illegal, [online intermediaries] might rightly be held liable for permitting malicious behaviors that they could have detected or deterred at reasonable cost.<sup>194</sup>

These cases and doctrines do not offer precise examples of how offline intermediary liability should translate to the online context, but they do offer several valuable tools. Most importantly, they demonstrate that the law has long wrestled with how to frame the legal duties owed by a service provider to its customers and the public, while also policing the bad acts of third parties. Additionally, although no example provides a perfect fit for the facts of online intermediary liability, these cases demonstrate how to think through the *principles* of holding intermediaries to a duty-of-care standard. Finally, while it was once in fashion to proclaim the Internet a wholly unique invention to which traditional laws could not readily be applied,<sup>195</sup> a more sober analysis of the history of the common law demonstrates that new business models and new technologies are regularly and inevitably incorporated into the law.

#### **IV. A Path Forward for Section 230 Reform**

No legal regime is perfect, particularly when, as in the case of Section 230, it was crafted specifically to govern a largely nascent industry that has subsequently grown by orders of magnitude in both size and importance. After more than two decades, it is eminently reasonable to reflect upon our accumulated experience with online platforms and to use that earned wisdom to refine the legal rules that govern platform liability. Reform should proceed in a way that adequately incorporates legitimate objections, however, many of which raise the important concerns that reform will have the unintended consequence of destroying vibrant free expression online and impairing the online economy.<sup>196</sup> Effective reforms that deter tortious and illegal content must also contain procedural safeguards that prevent an explosion of unmeritorious litigation. In short, reforms must be designed to pass a cost-benefit test that suggests that the likely benefits of changing the legal regime will exceed the likely costs.<sup>197</sup>

Toward that end, in this section we propose and discuss a set of reforms that, we believe, effectively balance these costs and benefits.

---

<sup>194</sup> Lichtman & Posner, *supra* note 24, at 256.

<sup>195</sup> See, e.g., John Perry Barlow, *A Declaration of the Independence of Cyberspace* (1996), available at <https://www.eff.org/cyberspace-independence>; Esther Dyson, George Gilder, George Keyworth, & Alvin Toffler, *Cyberspace and the American Dream: A Magna Carta for the Knowledge Age* (Progress & Freedom Foundation Future Insight No. 1.2, Aug. 1994), available at <http://www.pff.org/issues-pubs/futureinsights/fi1.2magnacarta.html>.

<sup>196</sup> See generally Part II, *supra*.

<sup>197</sup> See *Liability for User-Generated Content Online: Principles for Lawmakers*, *supra* note 14.

### A. Establishing a duty of care that balances immunity and accountability

There is little (if any) justification for imposing upon online platforms obligations that go *further* than the common law would require. But the common law has developed several standards of care for intermediaries in situations where the intermediary either otherwise prevents or reduces the direct enforcement of the law, or else where the intermediary is the least-cost avoider of harm, such that imposing upon it a duty of care results in the efficient level of precautions and activity to mitigate harm.

As noted, there is no perfect analogue in the common law to the relationships among online intermediaries, malfeasant users, harmed users, and harmed non-users.<sup>198</sup> But there are several analogous situations in the common law that offer guidance as to what a common law of online intermediary duty of care might encompass. Ultimately, however, it is a feature, not a bug, of our proposed regime that the courts will have to discover over time what such a duty of care should entail.

**First and foremost, we believe that Section 230(c)(1)'s intermediary-liability protections for illegal or tortious conduct by third parties can and should be conditioned on taking reasonable steps to curb such conduct, subject to procedural constraints that will prevent a tide of unmeritorious litigation.<sup>199</sup>**

Such a standard would do no more than impose upon online intermediaries the same sorts of obligations that the law imposes upon analogous offline intermediaries. Indeed, it would do quite a bit *less*, insofar as (as discussed below<sup>200</sup>) we propose that the law maintain a safe harbor even from most *litigation costs* for platforms that meet the standard, unlike for offline intermediaries. By the same token, of course, platforms that fail to meet this standard would lose the benefits of immunity.

---

<sup>198</sup> See *supra* Part III.C.

<sup>199</sup> This proposal is not entirely new, and others—notably Neil Fried and Danielle Citron—have been proposing some form of it for the last several years. See, e.g., *Disinformation Online and a Country in Crisis: Hearing before H. Subcomm. on Commc'ns. & Tech, and H. Subcomm. on Consumer Prot. and Commerce, H. Comm. on Energy and Commerce, 116th Cong. (2020)* (statement of Neil Fried, Principal, DigitalFrontiers Advocacy), available at <https://digitalfrontiersadvocacy.com/6-24-20-sec-230-testimony>; *Fostering a Healthier Internet to Protect Consumers: Hearing before H. Subcomm. on Commc'ns. & Tech, and H. Subcomm. on Consumer Prot. and Commerce, H. Comm. on Energy and Commerce, 116th Cong. (2019)* (statement of Prof. Danielle K. Citron, Boston University School of Law), available at <https://energycommerce.house.gov/committee-activity/hearings/hearing-on-fostering-a-healthier-internet-to-protect-consumers>. See also Citron & Wittes, *supra* note 46; Citron & Jurecic, *supra* note 86; Citron & Franks, *supra* note 37; Tom Wheeler, et al., *New Digital Realities; New Oversight Solutions in the U.S.*, Shorenstein Center on Media, Politics and Public Policy, Policy Paper (Aug. 2020), available at <https://shorensteincenter.org/new-digital-realities-tom-wheeler-phil-verveer-gene-kimmelman/> (supporting restoration of the duty of care but proposing to do so by creating an entirely new agency to regulate platforms rather than by simply narrowing section 230 immunity).

<sup>200</sup> See *infra* Part IV.C.



This does not mean that our proposal is insensitive to the relevant differences between the online and offline environments, and particularly to the differences between online and offline providers of third-party content. But, ultimately, the imposition of intermediary-liability law should be directed at better aligning platforms' incentives with the aims of existing statutory and common law such that they are quicker and more effective in preventing, finding, and removing illegal or tortious content. "To serve as an *ex ante* enforcement strategy... collateral liability... must prescribe a mechanism—an enforceable duty—that allows private parties to avert misconduct when they detect it."<sup>201</sup>

Thus, we note that limiting immunity is not the same thing as imposing liability. We should not hold online platforms vicariously liable for the speech of third parties, as we do offline publishers, because the volume of user-generated content online is so much greater than offline; because the relationship between online platforms and users is much more attenuated than that between, e.g., a newspaper and the authors of letters it chooses to publish; and because there are potentially large costs to overly chilling free expression online. Our proposal doesn't contemplate suits against platforms for the underlying tortious conduct of their users. Rather, we examine the more limited question of whether a platform took appropriate steps to prevent harm. Our proposal is thus sensitive to the high cost of over-deterrence in this context.<sup>202</sup>

Further, and for much the same reasons, we believe that additional safeguards are advisable. Thus, while we support subjecting online platforms to a duty of care with respect to their treatment of user-generated content despite the increased litigation risk, we believe that certain procedural protections are needed to mitigate the more substantial duck-bites problem that the scale of user-generated content online presents.<sup>203</sup> In this way our proposal diverges from, for example, Citron and Wittes' proposed "reasonableness" amendment to Section 230, which provides that:

No provider or user of an interactive computer service that takes *reasonable steps to prevent or address unlawful uses of its services* shall be treated as the publisher or speaker of any information provided by another information content provider *in any action arising out of the publication of content provided by that information content provider*.<sup>204</sup>

Although a move in the right direction, this proposal too greatly discounts the potential harms that could arise if, for instance, this standard resulted in throwing every trivial online platform dispute

---

<sup>201</sup> Kraakman, *Gatekeepers*, *supra* note 6, at 57.

<sup>202</sup> "Negligence-based standards serve as the traditional remedy for overdeterrence problems associated with 'positive externalities.'" Assaf Hamdani, *Who's Liable for Cyberwrongs?*, 87 CORNELL L. REV. 901, 934 (2002) (citing RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 729-37 (5<sup>th</sup> ed. 1998) (endorsing the adoption of negligence standards for libel, among other harmful forms of speech, because of the external benefits of speech)).

<sup>203</sup> See *infra* Part IV.C.

<sup>204</sup> Citron & Wittes, *The Internet Will Not Break*, *supra* note 46, at 419.

into a litigation process. Instead, our proposal reflects Reiner Kraakman’s admonition that informed criteria should guide courts’ adjudication of due-care obligations:

Without more, duties of due care or reasonable investigation are mere placeholders that shift the problem of duty design elsewhere. In the first instance, they shift the problem to the discretion of gatekeepers, subject to ex post review by the courts. But to the extent that this alone still leaves diffuse and potentially costly duties, *due care obligations require additional focusing devices to limit liability*. One of the best focusing devices is a community of knowledgeable gatekeepers who can tell courts or administrators what “due care” ought to mean by developing informed criteria of their own. After this, the next best alternative is to deduce the gatekeeper’s monitoring capabilities from established business practices. Widely shared business practices can help to focus gatekeeping duties....<sup>205</sup>

At the same time, compared to traditional media providers operating offline, online platforms host more behavior and commerce in general, and more that isn’t purely expressive; tortious and illegal content online are less susceptible to normal deterrence; and online content is disseminated both faster and more broadly. The risks and costs of unlawful conduct may thus be greater. The current Section 230 doesn’t just reduce the liability risk of intermediaries for user-generated content; it removes it virtually entirely. As we have discussed, that outcome seems insufficiently insensitive to the heightened threat of harm online.

Navigating the line between immunity and accountability for online intermediary liability—in accordance with the analysis above—suggests other principles, as well. One of the most important among these is the need for more nuanced treatment of speech and conduct—communication torts and non-communication torts—online.

---

<sup>205</sup> Kraakman, *Gatekeepers*, *supra* note 6, at 79-80 (emphasis added). Some critics may assert that this could lead to anticompetitive effects due to reliance on the practices of established social media companies. See, e.g., Armijo, *supra* note 141, at 12. But the standard of care for negligence has always been informed by custom, including industry standards, while allowing juries to determine whether a deviation from custom is unreasonable or even whether the custom itself is reasonable. See Kenneth S. Abraham, *Custom, Noncustomary Practice, and Negligence*, 109 COLUM. L. REV. 1784, 1786 (2009) (“The new Restatement of Torts, for example, sets it out with admirable clarity. Evidence of an actor’s compliance with custom is admissible for use defensively (as a ‘shield’) to show reasonable care, and evidence of an actor’s departure from custom is admissible for use offensively (as a ‘sword’) to show negligence. But neither form of evidence is conclusive. The finder of fact may determine that an actor who complied with custom was negligent, or that an actor who departed from custom exercised reasonable care.”). The burden should be upon those who wish to depart from the normal operation of tort law principles to show why *this* market is so different that traditional principles shouldn’t apply. Those defending the Section 230 status quo have failed to explain why relying on custom (i.e., industry standards) to establish a standard of care for content moderation by online intermediaries would have any greater anticompetitive effects than does relying on custom in any other industry. Cf. Armijo, *supra* note 141, at 12-13.

Thus, as an exception to the general reasonableness rule above, our proposal suggests maintaining Section 230(c)(1)'s categorical exemption from treating online platforms as speakers or as publishers (in the sense of defamation law's term of art<sup>206</sup>) *for communication torts*, and adopting a knowledge requirement for such claims. Put differently, online platforms should not face liability for communication torts arising out of user-generated content unless they fail to remove content they knew or should have known was defamatory.

That is to say, so long as a platform doesn't have actual knowledge (e.g., through receipt of a court order<sup>207</sup>) of defamatory content or doesn't fail to investigate when it has reason to believe that a piece of content is defamatory, it shouldn't be treated as a publisher of that content. Once it has such knowledge, however, it should have an obligation to make reasonable efforts to remove and prevent republication of the defamatory material. This is an extension of the common law rule for offline distributors of tortious content, keyed (again) to the relevant distinctions between offline and online intermediaries cutting both for and against heightened liability.<sup>208</sup>

The reason for maintaining intermediary immunity for communication torts to a greater degree than non-communication torts is twofold. First, as discussed above, it is apparently more in line with the original intent of Section 230 immunity, which aimed to mitigate the considerable online moderator's dilemma in the face of the threat of defamation liability.<sup>209</sup> Second, as mentioned above, there is an important, if imperfect, distinction between speech and conduct.<sup>210</sup> Activity that would be considered conduct if engaged in offline shouldn't automatically be considered speech when online.

While there are circumstances in which the distinction may be unclear, courts long have been able to distinguish speech from conduct over the course of considerable First Amendment jurisprudence.<sup>211</sup> Our proposal relies on this ability of courts to properly apply the standard for

---

<sup>206</sup> See *supra* note 40.

<sup>207</sup> We hasten to note that we can conceive of other forms of notice that could give rise to the knowledge requirement, but that, say, a single complaint by an offended user is unlikely to be sufficient. It is up to the courts to determine where to draw the line, however.

<sup>208</sup> See, e.g., *Lewis v. Time, Inc.*, 83 F.R.D. 455, 465 (E.D. Cal. 1979) (*aff'd* 710 F.2d 549 (9th Cir. 1983)) ("specific allegations concerning actual knowledge or giving rise to a duty to investigate are required in any lawsuit seeking to impose liability on a distributor for libelous material in a periodical over which he has no editorial control."). See also *Frye*, *supra* note 116 ("[T]he republication rule doesn't attribute statements to distributors without knowledge, and doesn't require immediate removal. Typically, ISPs voluntarily remove libelous statements, once they become aware of them. But it is unclear whether Section 230 shields ISPs from injunctions to remove libelous material. Some courts have held it does, and others have held it doesn't. Under the republication rule, ISPs would surely be liable for continuing to disseminate a third-party statement once they know it is libelous. I find it hard to believe that courts will ultimately construe Section 230 differently. At some point, refusal to remove a libelous statement must become an endorsement.").

<sup>209</sup> See *supra* notes 33-46 and accompanying text.

<sup>210</sup> See *supra* notes 56-69 and accompanying text.

<sup>211</sup> See *Citron & Franks*, *supra* note 37, at 58-60.

immunity. In the abstract, this would introduce additional complexity and expected litigation costs. But the payoff for this additional complexity is a simplified legal rule for communication torts that provides them even more protection than offline publishers receive. There is no reason to think that the net effect would be an increase in expected litigation costs.

A selection of key cases in which defendants have sought Section 230 immunity can help to illustrate the workability of the distinction. For instance, Silk Road founder Ross Ulbricht attempted to use a Section 230 defense to avoid liability for the dark-web site where users could buy illegal products and services using the tools of cryptocurrency and Tor.<sup>212</sup> No one could set up a platform to do this legally in the offline world and, despite the attempted Section 230 defense,<sup>213</sup> the court decided in Ulbricht's case that it couldn't be done in the online world either. Nor did the court accept the claim that Ulbricht's conduct was inherently speech and thus protected: "[T]he law is clear that speech which is part of a crime is not somehow immunized. For instance, no one would doubt that a bank robber's statement to a teller—'This is a stick up'—is not protected speech."<sup>214</sup> In other words, just because Silk Road users were engaged in online conduct didn't transform the illegal transactions into protected speech. Similarly, cases like *Oberdorf*,<sup>215</sup> *Lemon v. Snap*,<sup>216</sup> *Herrick v. Grindr*,<sup>217</sup> and *Armslist*<sup>218</sup> aren't really about speech, but about whether those platforms were designed in ways that facilitated tortious or illegal conduct. As directed by this proposal, courts will have to determine whether future causes of action are communications-related or about conduct.

At the same time, introducing a knowledge requirement, of course, implicates the moderators' dilemma, and the risk that the act of moderation will give rise to knowledge sufficient to impose potential liability.

**Thus, we propose that Section 230(c)(2)'s safe harbor should remain in force and that, unlike for traditional media operating offline, the act of reasonable content moderation by online platforms should not, itself, create liability exposure.**

---

<sup>212</sup> See, e.g., Nick Gillespie, *The Most Important Trial in America*, THE DAILY BEAST (Jan. 14, 2015), <https://www.thedailybeast.com/the-most-important-trial-in-america>.

<sup>213</sup> Though it was rejected because Section 230 immunity doesn't stop prosecutions for federal crimes. See 47 U.S.C. §230(e)(1). It would be an interesting question whether a state could have pursued criminal violations against the site or if Section 230 immunity would have protected Silk Road and Ulbricht from any claims based upon user conduct. See, e.g., Caleb Garling, *Contract killing aside, Silk Road may have been legal*, SAN FRANCISCO CHRON. (Oct. 9, 2013), <https://www.sfgate.com/technology/article/Contract-killing-aside-Silk-Road-may-have-been-4882819.php>.

<sup>214</sup> *United States v. Ulbricht*, 31 F. Supp. 3d 540, 568 (S.D.N.Y. 2014).

<sup>215</sup> *Oberdorf v. Amazon*, 930 F.3d 136 (3rd Cir. 2019).

<sup>216</sup> See *Lemmon v. Snap, Inc.*, Slip Op., No. 20-55295 (9th Cir. May 4, 2021).

<sup>217</sup> *Herrick v. Grindr, LLC*, 306 F. Supp. 579 (S.D. N.Y. 2018).

<sup>218</sup> *Daniel v. Armslist, LLC*, 926 N.W.2d 710 (Wis. 2019).

That is, the act of content moderation does not inherently give rise to constructive knowledge of illegal behavior. As the drafters of Section 230 were acutely aware, a presumption that a platform's decision to moderate demonstrated that it had constructive knowledge of illegal behavior would dramatically deter platforms from engaging in beneficial content moderation. While taking something down may, depending on the circumstances, suggest the platform knows it is unlawful, that act of moderation alone does not establish that the platform should necessarily be liable for failing to remove the same or similar content everywhere else it may appear on the platform.

It is important here not to conflate the lawfulness of the content with the reasonableness of the mechanisms used to find and remove it, although the two are not totally distinct. If removing one piece of content because it is illegal reasonably gives rise to an assumption that it exists elsewhere and would cause harm, then it could well be unreasonable not to attempt to remove other instances of the same content. There is not necessarily any reason to think this would be a common state of affairs, but if removing one piece of content in accordance with reasonable practices, in fact, makes it easier and less costly to find and remove other harmful instances of the same content, then it *should* create a greater obligation to do so; red-flag knowledge of illegal or tortious conduct should trigger *some* moderation responsibilities.

To summarize:

**Online platforms should operate under a duty of care obligating them to adopt reasonable content-moderation practices with respect to illegal or tortious third-party content. Section 230 should operate to prevent liability when the platform fails to moderate defamation or other communication torts and has no knowledge (actual or constructive) of the problem. And an act of moderation should not give rise to liability unless a failure to further moderate is unreasonable.**

In short, Section 230 should provide a safe harbor from liability when a platform takes reasonable steps to moderate unlawful conduct, and online intermediaries should not generally be held vicariously liable for the bad acts of third parties. But they should be held responsible for their own failures to comport with concrete obligations under a reasonable duty of care when they are the least-cost avoiders of harm. And, as we discuss below, there are some circumstances where aligning incentives may even suggest vicarious liability.<sup>219</sup>

This approach is hardly novel. As illustrated above,<sup>220</sup> the law has long imposed obligations on intermediaries, either through direct commands or as an inevitable consequence of their seeking to avoid liability. The threat of liability that intermediaries face will not always be for the underlying

---

<sup>219</sup> In general, this would arise when the platform itself is the source of the difficulty in enforcing laws or bringing private actions against offending content creators themselves. This circumstance is discussed below at notes 229-235 and accompanying text.

<sup>220</sup> See *supra* Part III.C.

conduct. Thus, for example, bars check patrons' identification not, in the first instance, because they might be held vicariously liable for a patron's underage drinking, but in order to avoid liability for serving alcohol to minors. This is a clear application of the "least-cost-avoider" principle. The ultimate objective is to curb underage drinking, and imposing liability on bars for serving alcohol to minors effectively "deputizes" bar owners to police underage drinking in a far more cost-effective manner than could be accomplished by the local police force acting on its own. Indeed, similar to our proposal, bars can typically avoid liability for serving alcohol to minors by pleading as an affirmative defense that they have properly inspected identification.<sup>221</sup> This device, then, acts as a safe harbor to induce bar owners to adopt cost-effective deterrent tactics without imposing an overly costly risk of liability on the act of serving alcohol generally. Similarly, the Digital Millennium Copyright Act adopts the same combination of liability and safe harbor: "Thanks to [the DMCA's safe harbor provision], Internet service providers and other firms associated with the Internet know that they are immune from indirect liability as long as they follow the guidelines explicitly set forth. This safe harbor thus eliminates the risk created by an otherwise uncertain legal standard."<sup>222</sup>

## **B. The outer boundaries of a duty of care**

"Progress lay[s] in the transition from requiring 'reasonable' conduct to defining what 'reasonableness' consist[s] of."<sup>223</sup> Developing a duty-of-care standard is not a simple task, to be sure, as there are no exact precedents in the common law for online platform liability. It may be the case that, in some circumstances, platforms would owe very little duty, while in others they may have heightened obligations.

But, at the margin, limiting immunity will magnify platforms' risks and costs, both of which are signals to the platforms to design their services in ways that mitigate their exposure to such risks and costs. This is how a common law duty of care works in other contexts. A modification to Section 230 that increases expected costs might be in order where those costs lead to legitimate findings of liability. There is a cost, as well, to *not* deterring unlawful conduct, which is discounted entirely under the current prevailing interpretation of Section 230. The role of tort law in internalizing the externalities created by online platforms is today foreclosed by Section 230 immunity.

---

<sup>221</sup> See, e.g., Jose Rivera, *Selling Alcohol to Minors*, LEGAL MATCH (Jun. 26, 2019), <https://www.legalmatch.com/law-library/article/selling-alcohol-to-minors.html> ("In most states, a licensee is only held liable if they sold to a minor without asking for any ID. If an ID card is asked for, and a fake ID indicating the minor is actually 21 is shown, then in almost all cases, no charges will be filed from either the police or from the state alcohol control board.").

<sup>222</sup> Landes & Lichtman, *supra* note 6, at 406. Note, we are not advocating the DMCA as a perfect model for intermediary liability. As with Section 230, Section 512 of the DMCA has shown its age over the years and could do with considerable reform. We merely point to it as an example of a workable, if highly imperfect, model in order to demonstrate that properly constrained liability rules can impose obligations on online providers without unavoidably catastrophic consequences.

<sup>223</sup> Frank H. Easterbrook, *Discovery as Abuse*, 69 B.U. L. REV. 635, 640 (1989) (citing OLIVER WENDELL HOLMES, *THE COMMON LAW* 111-13, 123-26 (1881)).

By the same token, we also do not want to deter beneficial conduct or impose excessive costs relative to benefits. Thus, reforms that open the door to potential liability under a duty-of-care standard must be mindful that a large volume of unmeritorious litigation could impose unreasonable costs that overwhelm the benefits of legitimate litigation.

A properly designed duty-of-care standard, moreover, should be flexible and account for the scale of a platform, the nature and size of its user base, and the costs of compliance, among other considerations. Indeed, this sort of flexibility is a benefit of adopting a “reasonableness” standard, such as is found in common law negligence.<sup>224</sup>

The optimal result is virtually never a maximum or minimum, practically speaking. No one wants rules that would guarantee zero harmful content, precisely because such rules would dramatically chill *legal* speech and conduct. Instead, as a society, we want laws that encourage as much expression and conduct and investment and innovation as possible, up to the point that the costs of *unlawful* expression and conduct, and/or the innovation and investment to facilitate the *unlawful* speech or conduct encouraged by such laws, outweigh the benefits of the legal speech and activity they engender.

By the same token, platforms should have discretion to set their own standards. But those standards cannot willfully ignore unlawful conduct such that the production of such content increases, any more than we would allow a shopping mall to willfully ignore how its lack of outdoor lighting or otherwise negligent security leads to increased thefts or assaults against customers in its parking lot at night.<sup>225</sup> Consequently, without more, one cannot “over-censor” unlawful conduct, as

---

<sup>224</sup> While some have argued that a negligence standard would prove unworkable in practice as applied to online platforms, see, e.g., Armijo, *supra* note 141, at 13-16, this proposal is specifically locating reasonableness as a standard of care for purposes of determining Section 230 immunity. It would be contingent upon the underlying causes of action (of which negligence theories are just one type) to determine to whom, if any, the online intermediaries owed a duty of care at all for the purposes of liability. Thus, insofar as critics are correct that intermediary liability would be quite limited under a negligent republication standard (as nonfeasance rather than misfeasance, see *id.* at 14-15), this shows how limited the effects of this proposal would be in encouraging lawsuits based on the speech torts Section 230 was passed in response to, and ironically, how little this proposal would chill speech. Cf. *id.* at 16-17.

<sup>225</sup> See, e.g., *Prime Hospitality Corp. v. Simms*, 700 So.2d 167 (1997) (upholding a jury verdict in favor of plaintiffs alleging negligent security at a hotel which was the proximate cause of a robbery and rape and awarding \$400,000 in damages). See also *The Value of Inadequate Security Lawsuits*, NEGLIGENT SECURITY ATTORNEY (last accessed Sept. 6, 2021), <https://www.negligentsecurityattorney.com/verdicts-settlements> (collecting many examples of malls, hotels, stores, apartment complexes, and restaurants being held liable for negligent security). *A fortiori*, of course, it is also not permitted to *knowingly* set up a community for the purpose of breaking the law. See *U.S. v. Ulbricht*, 31 F. Supp. 3d at 556 (“Silk Road was specifically and intentionally designed for the purpose of facilitating unlawful transactions. The Indictment does not allege that Ulbricht is criminally liable simply because he is alleged to have launched a website that was—unknown to and unplanned by him—used for illicit transactions. If that were ultimately the case, he would lack the *mens rea* for criminal liability. Rather, Ulbricht is alleged to have knowingly and intentionally constructed and operated an expansive black market for selling and purchasing narcotics and malicious software and for laundering money. This separates Ulbricht’s alleged conduct from the mass of others whose websites may—without their planning or expectation—be used for unlawful purposes.”).

users do not have a right to behave unlawfully.<sup>226</sup> Rather, we care about “over-censorship” of unlawful conduct only to the extent that, in seeking to deter or remove unlawful conduct, some quantity of lawful conduct is likely also to be deterred or removed, the loss of which may outweigh the value of the unlawful conduct deterred or removed.<sup>227</sup>

The approach we suggest would create some obligation for platforms to find and remove unlawful content. This could include a “stay-down” regime in which platforms face a heightened obligation not only to remove unlawful content when notified of it, but also to implement reasonable measures to try to prevent it from reappearing.<sup>228</sup> But any such obligations should be reasonable, taking account of the platform’s specific characteristics, reasonable technology, costs, etc.

But aligning the incentives of platforms with the law does suggest imposing a heightened standard in some contexts. When the platform itself is effectively responsible for impeding the “normal” operation of the legal regime, obligating it to act with reasonable care may prove impossible. In other words, the “reasonable” solution may be to prevent the platform from operating at all, which would unreasonably chill legal and desirable content.

---

<sup>226</sup> It is important to again note that our reasonableness proposal doesn’t change the fact that the underlying elements in any cause of action still need to be proven. It is those underlying laws, whether civil or criminal, that would possibly hold intermediaries liable without Section 230 immunity. Thus, for example, those who complain that FOSTA/SESTA harmed sex workers by foreclosing a safe way for them to transact (illegal) business should really be focused on the underlying laws that make sex work illegal, not the exception to Section 230 immunity that FOSTA/SESTA represents. By the same token, those who assert that Section 230 improperly immunizes “conservative bias” or “misinformation” fail to recognize that, because neither of those is actually illegal (nor could they be under current First Amendment law), Section 230 offers no additional immunity from liability for such conduct: There is no underlying liability from which to provide immunity in the first place.

<sup>227</sup> As represented in the doctrine of overbreadth and chilling effects. See Parker & Hudson, *supra* note 13; Askin, *supra* note 13. The Supreme Court’s jurisprudence on obscenity law (e.g., how “community standards” are applied) and defamation (e.g., causes of action by public figures requiring a showing of “actual malice”) also illustrate how substantive law has been limited by the First Amendment, representing the same idea that government regulation of unprotected speech is still limited if it sweeps in too much protected speech. See *Ashcroft v. ACLU*, 535 U.S. 564, 586-89 (2002) (O’Connor, J., concurring); *New York Times v. Sullivan*, 376 U.S. 254 (1964).

For our purposes, the applicability of underlying laws without Section 230 immunity is limited by the elements of those laws as well as the First Amendment. See Armijo, *supra* note 141, at 24 (discussing *Smith v. California*, 316 U.S. 147 (1959) and how it limited the ability to hold booksellers liable for carrying obscene literature); *id.* at 26-33 (discussing the (in)applicability of incitement laws to online platforms under an intermediary-liability theory). In other words, the lack of “reasonable content moderation” for the purposes of this proposal wouldn’t automatically make online intermediaries liable if there is no constitutionally valid law that could apply.

<sup>228</sup> Of course, this is what many platforms already do today to keep objectionable content from immediately reappearing online. While these efforts are necessarily imperfect, reasonableness does not require perfection. Perhaps the most well-known example of both this practice as well as its inherent imperfection is Facebook’s effort to keep the Christchurch shooting video from constantly reappearing following its initial removal. See Kate Klonick, *Inside the Team at Facebook That Dealt with the Christchurch Shooting*, THE NEW YORKER (Apr. 25, 2019), available at <https://www.newyorker.com/news/news-desk/inside-the-team-at-facebook-that-dealt-with-the-christchurch-shooting>.



**Thus, when a platform operates in a fashion that impedes the application of direct liability for user-generated content, it should be offered a choice: risk vicarious liability if a court finds that the law is broken (or a tort is caused) by content it hosts, or else mitigate whatever dynamic it has created that impedes direct law enforcement.**

The most obvious circumstance in which this would arise concerns the ability of courts to exercise jurisdiction over end users when a platform's normal operation renders those users anonymous. Thus, one component of an improved liability regime for online intermediaries should be an obligation that intermediaries unmask anonymous users in certain situations (or, alternatively, operate without anonymity), consistent with the characteristics and capacities of the platform.

Writing about the holding in *Oberdorf v. Amazon*,<sup>229</sup> which provides the best example of a court's effort to implement such an approach, Gus Hurwitz notes:

Section 230's immunity could be attenuated by an obligation to facilitate the identification of users on that platform, subject to legal process, in proportion to the size and resources available to the platform, the technological feasibility of such identification, the foreseeability of the platform being used to facilitate harmful speech or conduct, and the expected importance (as defined from a First Amendment perspective) of speech on that platform.

In other words, if there are readily available ways to establish some form of identify for users... and there is reason to expect that users of the platform could be subject to suit... then the platform needs to be reasonably able to provide reasonable information about speakers subject to legal action in order to avail itself of any Section 230 defense. Stated otherwise, platforms need to be able to reasonably comply with so-called unmasking subpoenas issued in the civil context to the extent such compliance is feasible for the platform's size, sophistication, resources, etc.<sup>230</sup>

Such a reform should be unobjectionable; Section 230 was never contemplated as a liability shield for content creators, as opposed to intermediaries.<sup>231</sup> Further, such a reform is consistent with imposing upon platforms a reasonableness requirement where the platform is the least-cost avoider. Platforms "need to make reasonable efforts to ensure that their users engaging in problematic speech

---

<sup>229</sup> *Oberdorf*, 930 F.3d at 136.

<sup>230</sup> Gus Hurwitz, *The Third Circuit's Oberdorf v. Amazon Opinion Offers a Good Approach to Reining in the Worst Abuses of Section 230*, TRUTH ON THE MARKET (Jul. 15, 2019), <https://truthonthemarket.com/2019/07/15/the-third-circuits-oberdorf-v-amazon-opinion-offers-a-good-approach-to-reining-in-the-worst-abuses-of-section-230/>.

<sup>231</sup> As Hurwitz points out, "[i]n an era in which sites like 8chan expressly don't maintain user logs in order to shield themselves from known harmful speech, and Amazon Marketplace allows sellers into the market who cannot be sued by injured consumers, this is a common-sense change to the law." *Id.* See also *id.* ("The proposal offered here is not that platforms be able to identify their speaker—it's better described as that they not deliberately act as a liability shield. It's [sic] requirement is that platforms implement reasonable identity technology in proportion to their size, sophistication, and the likelihood of harmful speech on their platforms.... This would restore the status quo ante, under which intermediaries and agents cannot be used as litigation shields without themselves assuming responsibility for any harmful conduct. This shielding effect was not an intended goal of Section 230, and it has been the cause of Section 230's worst abuses.").

can be identified by parties harmed by their speech or conduct.”<sup>232</sup> If they fail to do so, Section 230 immunity should not shield them from suit.

While Hurwitz views this reform as an alternative to a more generalized reasonableness standard, we believe it is an important adjunct that properly accounts for the problem of rapid and broad dissemination of tortious or illegal content online. Standing on its own, this reform addresses only the lowest-hanging fruit: those instances where a platform is not only the least-cost avoider but is, in fact, the *cause* of heightened enforcement costs. As such (and as noted) it should be received as unobjectionable.<sup>233</sup> But this leaves a significant number of situations unaddressed. Thus, to capture these while still acknowledging that a broader reasonableness regime imposes additional complications,<sup>234</sup> instead of limiting the possibility of liability to situations where a platform actively impedes normal judicial operation, we propose other procedural safeguards (e.g., a safe harbor for certified moderation practices) that would permit courts to develop broader reasonableness standards without unduly curtailing third-party content.

Importantly, such a requirement would still permit anonymous or pseudonymous speech or activity online. The aim is not to unmask users in the normal operation of the platform; rather, the aim is to facilitate the courts’ exercise of valid jurisdiction, limiting the unmasking of users only to the extent necessary for a valid case against an end user to proceed.<sup>235</sup>

By the same token, it could be sensible to impose upon online platforms an obligation to facilitate reporting illegal content not only to the platform itself, but to the proper legal authorities, in order to facilitate their taking direct action against the creators of illegal content. Currently, Section 230 has carve-outs for federal crimes, human trafficking, and child pornography. But the fact that intermediaries *may* be held liable for such content certainly doesn’t mean that *only* enforcement against platforms (as opposed to the actual content creators) is desirable or optimal.

---

<sup>232</sup> *Id.*

<sup>233</sup> This is not to say that, as with other aspects of our proposal, there would be no difficulty at all in implementing such a reform. But as Hurwitz aptly notes:

To be sure, there are still some uncomfortable and difficult substantive questions—has a platform implemented reasonable identification technologies, is the speech on the platform of the sort that would be viewed as requiring (or otherwise justifying protection of the speaker’s) anonymity, and the like. But these are questions of a type that courts are accustomed to, if somewhat uncomfortable with, addressing. They are, for instance, the sort of issues that courts address in the context of civil unmasking subpoenas.

*Id.*

<sup>234</sup> Hurwitz rejects the broader reasonableness approach because it “is problematic on both First Amendment and process grounds, because it requires courts to evaluate the substantive content and speech decisions that platforms engage in. It effectively tasks platforms with undertaking the task of the courts in developing a (potentially platform-specific) law of content moderations....” *Id.* As we discuss, our proposal should mitigate these problems, as well.

<sup>235</sup> We are aware that even this limited de-anonymization requirement could make some business models difficult to sustain. Arguably that is a reasonable cost, but it is one that should be weighed in evaluating such a proposal. And, of course, an online intermediary would always be free to offer completely anonymous services.

One way to operationalize this is to “[create] new reporting flows that would make it easier to hold people accountable and resolve disputes.”<sup>236</sup> Thus Matt Perault has suggested that:

[P]latforms could provide functionality that enables people to report content not only to the platform, but also to the offices of state attorneys general. Alternatively, platforms could provide options to report false voting information to an election-monitoring organization, to report harassment to victims’ support services, or to report defamation to lawyers who specialize in defamation law.<sup>237</sup>

### C. Additional procedural protections

Platforms should not bear excessive costs for conduct that does not and should not give rise to liability, while they should internalize the costs of responding to actual harms and meritorious litigation.

There are various ways that these goals can be accomplished, including reforms to civil procedure, relying on a regulatory agency to oversee creation of a duty of care, and implementing a “safe harbor” or presumption of reasonableness, among others. Likely, an optimal solution will contain some mixture of all these options.<sup>238</sup>

#### I. Establishing and updating the duty of care

The first hurdle that needs to be overcome is how to establish the duty of care. Had Section 230 never been enacted, courts likely would have continued to wrestle with cases in the early years of the Internet boom and, over time, would have established a reasonably consistent duty of care that applied to online intermediaries. It is certainly possible that such development could be thrown to common law courts today. But given the volume of economic and social activity that occurs online and the significant reliance interests in the status quo, such an approach could be greatly disruptive and ill-advised in the short term.

---

<sup>236</sup> Matt Perault, *Section 230 Reform: A Typology of Platform Power*, 2 CPI ANTITRUST CHRON. 14, 22 (May/Spring 2021).

<sup>237</sup> *Id.*

<sup>238</sup> Although we focus here on establishing a duty of care and “safe harbor” for compliance with it, other procedural reforms should be considered, including, e.g., limiting standing to government enforcers; limiting or prohibiting monetary penalties; and/or providing for fee-shifting. In particular, reforms that develop a duty-of-care model could also limit the parties to litigation in some manner. For example, enforcement could be limited to the FTC or to state attorneys general, with no private right of action available. It should be noted, however, that if the above notice-and-takedown system and adequate complaint-stage modifications (discussed below) are implemented correctly, it should not be necessary to limit the litigants. This is to say, if the platforms are reliably able to remove most illicit content, either proactively or in response to user notices, the remaining volume of litigation *should* be fairly restricted to meritorious claims. Further, coupled with complaint-stage civil procedure changes (discussed in the next section), the litigation that proceeds to discovery should be almost strictly composed of meritorious claims.

**Thus, we propose the establishment of “certified” moderation practices, compliance with which would operate to foreclose litigation at an early stage against online intermediaries in most circumstances.**

Specifically, we propose developing a hybrid approach that relies on statutory reform and common law courts, as well as industry standard-setting organizations and a multi-stakeholder body overseen by a federal agency like the Federal Trade Commission (FTC), the National Telecommunications and Information Administration (NTIA), the National Institute of Standards and Technology (NIST), the Cybersecurity and Infrastructure Security Agency (CISA), or some other regulatory body with expertise in law enforcement online.<sup>239</sup>

In order to receive the procedural benefits we propose in the next section, platform operators would be required to adopt a set of moderation practices, in accordance with industry-specific standards established by an industry standard-setting organization authorized by the FTC (or other agency), to deal with the potentially tortious or criminal use of their services by bad actors.

**The overseeing agency would be responsible for periodically convening a multi-stakeholder group in order to evaluate certified moderation practices, as well as answers that rely on certifications created during the pleading stage of litigation (discussed below). The multi-stakeholder group would produce a report that describes the latest best practices with which platform operators of various types should comply. Thus, when litigation does arise, platforms will be able to produce evidence of relevant best practices for a particular time period, and demonstrate how they did (or did not) comply with those practices.**

The overseeing agency would not itself be responsible for promulgating the standards, but for evaluating and approving their compliance with overarching standards established through the multi-stakeholder process.

**Importantly, we also believe that it is necessary that our proposal include a sunset provision that automatically terminates involvement of the federal agency and multistakeholder body after some time, and allows courts to continue developing the duty of care after a certain period of time.**

These elements differentiate our proposal from other facially similar proposals in critical ways. Tom Wheeler, Phil Verveer, and Gene Kimmelman, for example, have proposed the creation

---

<sup>239</sup> We have discussed elsewhere how the FTC has failed to develop what “reasonable” data security means in its Section 5 data security enforcement actions. See Geoffrey A. Manne & Kristian Stout, *When “Reasonable” Isn’t: The FTC’s Standardless Data Security Standard*, 15 J.L. ECON. & POL’Y 67 (2018). Part of the problem for the FTC is that it has abused its processes to strong-arm consent decrees, without defining what “reasonable” actually means and, as a practical matter, without submitting its consent orders to judicial oversight. Our proposal is different because it relies on private industry standards and feedback from the judiciary to establish what is reasonable, rather the agency itself. However, the FTC could be well-suited to enforcing the standards developed by private bodies more effectively under this proposal than it has been in independently developing what reasonable data security entails under its Section 5 authority.

of a “new cooperative industry-government regulatory model” that would rely on a “Code Council” comprised of members of industry and the public that is overseen by a new “Digital Platform Agency” (DPA).<sup>240</sup> A key difference between our proposals, however, is that the DPA would be responsible for both overseeing the creation of a code of conduct as well as enforcing its adherence among industry participants.<sup>241</sup> Our proposal, by contrast, relies on a federal agency only to kickstart and oversee a common law process, but ultimately contemplates ceding the entire process to the courts. As litigation proceeds based on these best practices, and as the multi-stakeholder body issues its periodic revisions, a common law-like body of principles will emerge and evolve to define and refine the duty of care that platforms owe their users.

Importantly, having an established duty of care contemplates that businesses may need to address harms proactively, and not simply react after harm has already occurred. The idea behind the duty of care is that businesses have an obligation to take reasonable steps to *prevent* harm. This ensures that platforms are answerable *ex post* when harm occurs, as well as *ex ante* by preventing harms under the established standards.

To facilitate this process at scale, some form of notice and takedown could be established—either by legislation, by courts as they address the reasonableness of platforms’ moderation practices, or *sua sponte* by platforms themselves as they respond to the development of legal standards—to help platforms discover tortious or illegal content and have a reasonable opportunity to remove it. Notably, a notice-and-takedown process would also provide the opportunity for users who have been “incorrectly” moderated to appeal to have their content restored. Such a system would also help to establish the conditions under which a platform would be deemed to be aware, or under which it should have been aware, of unlawful or tortious behavior on its service. It seems self-evident, for example, that a platform that is made aware of illegal user-generated content by a court order that has been shared with it can be deemed to have the requisite knowledge. But courts (or, conceivably, legislators) would surely establish circumstances short of service of a court order under which such actual or red-flag knowledge would be appropriate. A notice-and-takedown system would formalize these circumstances.

None of this mean the platform would automatically be held liable, of course, but it would lose the liability shield of Section 230 if a court were to determine that the evidence demonstrated that the certified moderation standards were not met. Moreover, even *if* a court permitted litigation to proceed beyond the complaint stage, a plaintiff would still have to prove that the platform was

---

<sup>240</sup> Tom Wheeler, Phil Verveer & Gene Kimmelman, *New Digital Realities; New Oversight Solutions in the U.S., The Case for a Digital Platform Agency and a New Approach to Regulatory Oversight*, Shorenstein Center Discussion Paper (Aug. 2020), available at [https://shorensteincenter.org/wp-content/uploads/2020/08/New-Digital-Realities\\_August-2020.pdf](https://shorensteincenter.org/wp-content/uploads/2020/08/New-Digital-Realities_August-2020.pdf)

<sup>241</sup> *See id.* at 20-21.

culpable for failing to meet its duty of care; that is, failure to adopt or comply with a certified standard would not result in automatic liability.

## 2. *Pleading practice changes*

Given the sheer volume of content online and the complexity, imprecision, and uncertainty of moderation processes, even very effective content-moderation algorithms will fail to prevent all actionable conduct, which could result in many potential claims. At the same time, it can be difficult to weed out unlawful conduct without inadvertently over-limiting lawful activity. A final refinement at the litigation stage is necessary to mitigate the risk of liability (and most litigation costs) when a platform demonstrates compliance with its certified best practices.

**Thus, in addition to the safe harbor of 230(c)(2), which ensures that platforms face no litigation risk for accidentally *over*-moderating content, Section 230 should also include a safe harbor that similarly removes litigation risk when, as is inevitable, platforms let injurious content slip through despite reasonable efforts. Our proposal would almost entirely foreclose further discovery and ongoing litigation against an online-platform defendant where it provides its certified moderation procedures in response to a complaint based on injury resulting from user-generated content.**

To the extent that the primary concern with frivolous litigation is plaintiffs' ability to impose excessive discovery costs on defendants, the availability of discovery should be sensibly limited in some manner short of foreclosing suit entirely (as is effectively the case under the current Section 230's immunity shield). In the typical case, a defendant faces a choice between answering a complaint or moving for dismissal under 12(b)(6). But there is no inherent reason that these must be mutually exclusive—indeed, there are circumstances in which answers and motions to dismiss are filed contemporaneously. Regardless of when it is filed, however, “an increase in the [motion-to-dismiss] grant probability will reduce the number of cases plaintiffs are willing to file.”<sup>242</sup> Thus, one important procedural limitation is directed at increasing the likelihood that appropriately answered complaints have a high likelihood of winning dismissal.<sup>243</sup> Not only will this greatly reduce the risk of excessive litigation costs, it will in most cases preclude the filing of a complaint in the first place.

As Judge Easterbrook notes, it is extremely difficult, but not impossible, to design a system to effectively internalize discovery costs. The key lies not in tinkering with limits on the number of interrogatories or in relying on the courts' ability to distinguish abusive discovery requests from appropriate ones: “[W]e cannot do anything about impositional requests that masquerade as proper

---

<sup>242</sup> Gelbach, *Locking the Doors to Discovery*, *supra* note 123, at 2305-06.

<sup>243</sup> Technically this would be a 12(c) motion for “judgment on the pleadings” rather than a 12(b)(6) motion to dismiss for “failure to state a claim upon which relief can be granted.” But both occur before “matters outside the pleadings are presented to and not excluded by the court” (i.e., discovery), after which a defendant moving for judgment would file a Rule 56 “motion for summary judgment.” *Cf.* Fed. R. Civ. P. 12(c); Fed. R. Civ. P. 12(b)(6). *See also* Fed. R. Civ. P. 56.

ones, and we have made it impossible to tell the two apart.”<sup>244</sup> Rather, what is required is clarity in the underlying proceedings:

[E]xcessive discovery is only a symptom of larger problems—the inability of our legal system to define what is relevant to a legal conflict and to make the parties bear the costs of their own endeavors. Relief comes from dealing with the causes....

...The principal facilitators of impositional discovery requests are rules (standards, really) that make everything relevant and nothing dispositive. Such approaches engender endless search for... well, for something that may turn out to be useful, once lawyers learn what the tribunal thinks important. If we want to cope with the ‘problem’ of discovery, we must do away with multi-factor standards, replacing them with rules that call for inquiry into a limited number of objectively ascertainable facts.<sup>245</sup>

To an important extent, *Twombly/Iqbal* mitigated this problem by requiring much greater concreteness in notice pleading—akin, in a sense, to what would be required by clearer underlying rules.<sup>246</sup> Even before *Twombly/Iqbal*, the Supreme Court had already begun demanding greater concreteness in analogous cases involving the immunity granted to government actors when that immunity was undermined by excessive discovery.<sup>247</sup> Although the source of and reasons for immunity in these cases is different, the court reiterated that qualified immunity, not absolute immunity, should be the norm. As the court noted, implicit in this policy conclusion is an “expectation that insubstantial suits need not proceed to trial.”<sup>248</sup> Ensuring proper freedom from vexatious lawsuits instead required “an adjustment of the ‘good faith’ standard established by [the court’s] decisions.”<sup>249</sup>

In short, the Court opted for a liability standard that turned on *objective* rather than *subjective* elements:

The subjective element of the good-faith defense frequently has proved incompatible with our admonition in *Butz* that insubstantial claims should not proceed to trial....

---

<sup>244</sup> Easterbrook, *Discovery as Abuse*, *supra* note 223, at 644.

<sup>245</sup> *Id.* at 643.

<sup>246</sup> See *supra* notes 121-125 and accompanying text.

<sup>247</sup> Easterbrook, *Discovery as Abuse*, *supra* note 223, at 643 (“Courts are beginning to do this. The Supreme Court recognized that its approach to official immunity led to endless rummaging, undercutting many of the goals of immunity doctrines. It responded by simplifying the rules and making cases turn on objective evidence.”).

<sup>248</sup> *Harlow v. Fitzgerald*, 457 U.S. 800, 808 (1982) (holding that presidential aides do not merit absolute immunity but do deserve certain protections from vexatious civil suits) (citing *Butz v. Economou*, 438 U. S. 478, 507-08 (1978)).

<sup>249</sup> *Id.* at 815.

...We therefore hold that government officials performing discretionary functions, generally are shielded from liability for civil damages insofar as their conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known.

Reliance on the objective reasonableness of an official's conduct, as measured by reference to clearly established law, should avoid excessive disruption of government and permit the resolution of many insubstantial claims on summary judgment. On summary judgment, the judge appropriately may determine, not only the currently applicable law, but whether that law was clearly established at the time an action occurred.<sup>250</sup>

Importantly, the Court's holding was premised on the need to deter inappropriate conduct and to compensate victims. The aim was to provide an objective basis for such a determination in order to minimize the burdens of unmeritorious suits, while still preserving the law's deterrent effect: "By defining the limits of qualified immunity essentially in objective terms, we provide no license to lawless conduct. *The public interest in deterrence of unlawful conduct and in compensation of victims remains protected by a test that focuses on the objective legal reasonableness of an official's acts.*"<sup>251</sup>

Our proposal pursues a similar course by converting, to the extent possible, what is inherently an indeterminate standard into a concrete and well-defined rule.<sup>252</sup> This is worth contrasting to the FTC's data-security enforcement practice under Section 5, which has failed to make "reasonableness" into anything concrete.<sup>253</sup> There, the FTC relies on enforcement actions to strong-arm companies into consent decrees, turning its reasonableness analysis into a de facto strict-liability standard.<sup>254</sup> This has led to a situation where the Eleventh Circuit held that an FTC

---

<sup>250</sup> *Id.* at 816-18.

<sup>251</sup> *Id.* at 819 (emphasis added), accord *Anderson v. Creighton*, 483 U.S. 635, 638-39 (1987) ("When government officials abuse their offices, 'action[s] for damages may offer the only realistic avenue for vindication of constitutional guarantees.' On the other hand, permitting damages suits against government officials can entail substantial social costs, including the risk that fear of personal monetary liability and harassing litigation will unduly inhibit officials in the discharge of their duties.... Somewhat more concretely, whether an official protected by qualified immunity may be held personally liable for an allegedly unlawful official action generally turns on the 'objective legal reasonableness' of the action, assessed in light of the legal rules that were 'clearly established' at the time it was taken.") (quoting *Harlow v. Fitzgerald*, 457 U.S., at 814, 818, 819) (emphasis added). Of course, as current disputes attest, even this qualified immunity appears in practice much closer to absolute immunity because the Court's objective standard is so rigidly applied that getting past summary judgment based on the defense of immunity is extremely difficult. See, e.g., Ben Sperry, *How Qualified Immunity Promotes the Unreasonable Use of Force by Police Officers*, THE HILL (Jun. 2, 2020), <https://thehill.com/blogs/congress-blog/politics/500673-how-qualified-immunity-promotes-the-unreasonable-use-of-force-by>; Ben Sperry, *Setting Up a Fair System for Determining Police Misconduct: Towards a Law & Economics Analysis of Qualified Immunity*, TRUTH ON THE MARKET (May 28, 2020), <https://truthonthemarket.com/2020/05/28/setting-up-a-fair-system-for-determining-police-misconduct-towards-a-law-economics-analysis-of-qualified-immunity/>.

<sup>252</sup> Consistent, in other words, with Reinier Kraakman's admonition that "[t]o serve as an ex ante enforcement strategy... collateral liability... must prescribe a mechanism—an enforceable duty—that allows private parties to avert misconduct when they detect it." Kraakman, *Gatekeepers*, *supra* note 6, at 57.

<sup>253</sup> See Manne & Stout, *supra* note 239.

<sup>254</sup> *Id.* ("In actuality, however, the Commission's manufactured 'reasonableness' standard—which, as its name suggests, purports to evaluate data security practices under a negligence-like framework—actually amounts in effect to a rule of strict liability for any company that collects personally identifiable data.")



injunction requiring “reasonable” data-security practices was so indeterminate as to be unenforceable.<sup>255</sup> For our proposal, however, the FTC would be in charge of the more modest task of convening and accepting standards developed by industry itself.

Our proposal does not, however, entirely foreclose the possibility of further discovery and litigation. The adoption of reasonable moderation procedures establishes a baseline adherence to the duty of care, and the implementation of our proposed “certified answer” procedure should ensure that costly litigation over reasonable, but inherently imperfect, moderation does not impose excessive costs on compliant online intermediaries. But adoption of such procedures does not necessarily imply their reasonable *implementation*. And, at the same time, novel and changing circumstances may undermine the reasonableness of even certified procedures. Thus, a process is needed to allow victims to proceed where online intermediaries have unreasonably failed to comply with their certified procedures, and in order for courts to evaluate certified practices and ensure that the judicial determination of reasonableness evolves with changing circumstances:

**In litigation, after a defendant answers a complaint with its certified moderation practices, the burden would shift to the plaintiff to adduce sufficient evidence to show that the certified standards were not actually adhered to. Such evidence should be more than mere *res ipsa loquitur*; it must be sufficient to demonstrate that the platform should have been aware of a harm or potential harm, that it had the opportunity to cure or prevent it, and that it failed to do so. Such a claim would need to meet a heightened pleading requirement, as for fraud, requiring particularity.<sup>256</sup>**

For example, a plaintiff could provide affidavits demonstrating that she had submitted notices of tortious content created by a particular user on multiple separate occasions, that the content or user was not dealt with as per published moderation standards, and that the plaintiff suffered a demonstrable harm at the hands of the user or as a result of the published content. If a plaintiff can meet this burden, the case could proceed without Section 230 immunity. In order to recover, however, the user would still ultimately have to prove that the platform actually violated its duty of care.

Moreover, even if such a case proceeded, platform operators would be liable only for failure to adhere to the standard, even if it turned out that, in a given case, the court determined that the standard was not, as applied to novel facts, reasonable. This sort of finding would provide the judicial feedback into the standard-setting process that enables the operation of common-law evolution. But the need for procedural limitations in this context means that such a finding would *not* be the basis for immediate liability; rather, the finding would represent an indication that current practices need

---

<sup>255</sup> LabMD, Inc. v. FTC, 894 F.3d 1221, 1300 (11th Cir. 2018).

<sup>256</sup> See Fed. R. Civ. P. 9(b) (“In alleging fraud or mistake, a party must state with particularity the circumstances constituting fraud or mistake.”).

to be adjusted to maintain their reasonableness.<sup>257</sup> Of course, collateral estoppel would ensure that a platform that fails to comply in the next case may be held liable.

In whichever way the complaint stage resolves, the platform operator would be required to submit the complaint, its answer, and the disposition of its motion to dismiss to the overseeing agency and multi-stakeholder body for use in their periodic review of moderation standards. This provides a practical feedback loop to help the agency and multi-stakeholder body better understand how the published standards work in practice.

Finally, another potential reform might be to consider limitations on the types of recovery available. For example, recovery might be limited to injunctions, assigning statutory damages for different kinds of harm, or permitting only the FTC to pursue monetary compensation as part of a Section 5 case under the moderation standards.

## V. Conclusion

Section 230 reform need not open the floodgates to every ill-conceived idea about the dangers of “Big Tech.” While there are certainly tradeoffs inherent in any regulatory regime—including the current Section 230 immunity regime—reform can be beneficial if the marginal benefits exceed costs. The approach advocated herein is designed to preserve the benefits of Section 230 immunity while allowing the law to hold online platforms accountable when they are in the best position to handle illegal conduct.

[Our view] admits room for the state to establish a framework of neutrally administered and enforced rules against which individuals arrange their private ordering.... Technological innovations do often offer significant benefits (not only in terms of liberty and autonomy, but general consumer welfare), of course, and any benefits arising from the adaptation and application of existing legal rules should be weighed against the possible costs of deterring the creation or welfare-enhancing deployment of technology. But, in principle, any technology,

---

<sup>257</sup> In this way, our proposed process is, again, notably different than the FTC’s data security enforcement process. That process does provide something of an analogy, insofar as it is an effort to impose a duty of care on companies that hold personally identifiable information to protect against harmful conduct undertaken by third parties (e.g., hackers). But as it has unfolded in practice, the FTC’s approach to data security regularly imposes liability for allegedly “unfair” data security practices that companies simply could not have known in advance would be insufficient. Although some refer to this FTC process as a “common law of data security”, see, e.g., Daniel Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583 (2014), it is actually anything but. See Geoffrey A. Manne & Ben Sperry, *FTC Process and the Misguided Notion of an FTC “Common Law” of Data Security*, ICLE Data Security & Privacy Working Paper (2014) at 13, available at <https://laweconcenter.org/resource/ftc-process-misguided-notion-ftc-common-law-data-security/> (“In this sense the FTC’s data security settlements aren’t an evolving common law—they are a static statement of ‘reasonable’ practices, repeated about 55 times over the years and applied to a wide enough array of circumstances that it is reasonable to assume that they apply to all circumstances. This is consistency. But it isn’t the common law.”). Our aim is to provide a somewhat related mechanism that does, in fact, incorporate meaningful procedural protections and the promise of a common-law, evolutionary standard.

no matter how revolutionary, can be brought within the ambit of predictable, neutrally administered legal rules.<sup>258</sup>

This approach creates no new causes of action; it does, however, require that online intermediaries seeking to avail themselves of Section 230 immunity must make reasonable efforts to deal with illegal activity on their platforms.

Critics of Section 230 reform are right to be concerned that changes to the law could threaten the large social gains the Internet economy has provided over the last quarter century. But such concerns ought not preclude adjustments to the legal regime that would better align platforms' incentives with the goal of mitigating the amount of damage caused by illegal and tortious conduct. Any changes must be considered as part of an analysis that weighs both the costs and benefits of a given reform, as well as the costs and benefits of the status quo. Too often, opponents of reform discount or ignore the costs of the current system; assume the benefits (like "more speech") are of infinite value; and/or ignore the benefits that may arise from a more optimized legal regime. Even modest reforms that directly address illegal conduct online could theoretically have large social benefits. Fears of censorship should be seriously assessed but should not be presumed to answer the question preemptively.

And to address the concern that open-ended litigation would arise with a more attenuated Section 230 liability shield, procedural reforms are likely necessary. As we discuss above, if properly implemented, there should be only modest, immediate effects on intermediaries. And as the duty of reasonable care is progressively interpreted by courts and regulators, we should expect experiments in new forms of content moderation, set against this backdrop of procedural safeguards. The result should be a progressive evolution toward ever-more-optimal practices.

Finally, this proposal does not demand perfection from platforms in their content-moderation decisions—only that they make reasonable efforts. The reasonableness standard is inherently flexible. What is appropriate for YouTube, Facebook, or Twitter will not be the same as what's appropriate for a startup social-media site, a web-infrastructure provider, or an e-commerce platform. Importantly (and among many other characteristics), courts could and should permit different practices depending on where an intermediary sits in the Internet stack. "Requirements drafted with user-facing services in mind will likely not work for these non-user-facing services."<sup>259</sup>

---

<sup>258</sup> Justin (Gus) Hurwitz & Geoffrey A. Manne, *Classical Liberalism and the Problem of Technological Change*, Working Paper (April 20, 2018) at 37, available at <https://ssrn.com/abstract=3384300> (subsequently published in *THE CAMBRIDGE HANDBOOK OF CLASSICAL LIBERAL THOUGHT* (M. Todd Henderson, ed. 2018)).

<sup>259</sup> *Liability for User-Generated Content Online*, *supra* note 14, at 2. The relevant Principle, in full, reads: "Principle #7: Section 230 should apply equally across a broad spectrum of online services. Section 230 applies to services that users never interact with directly. The further removed an Internet service—such as a DDOS protection provider or domain name registrar—is from an offending user's content or actions, the more blunt its tools to combat objectionable content become. Unlike social media companies or other user-facing services, infrastructure providers cannot take measures like removing individual posts

Also important is that courts could and should factor in the volume of posts on a user-facing platform; the costs of moderation; the resources available to the platform; and the benefits for and risks posed by users of its services. The effort needed to meet the reasonableness standard will inherently be proportional to platform size, ensuring that smaller platforms are not unreasonably burdened as they try to grow, and that firms are asked to expend resources only to the extent they make sense in the context of the severity and predictability of a potential harm. When a platform has fewer resources, less will be expected of it; when it has fewer users and uses, there will be less need to moderate. As the platform grows, so must its moderation efforts, but it will have more resources at its disposal. Allowing courts to apply the flexible common law duty of reasonable care would also enable the jurisprudence to evolve with the changing nature of online platforms, the problems they pose, and the moderating technologies that become available.

---

or comments. Instead, they can only shutter entire sites or services, thus risking significant collateral damage to inoffensive or harmless content. Requirements drafted with user-facing services in mind will likely not work for these non-user-facing services.”