RESEARCH PARTNER

Digital Planet | FLETCHER CENTER FOR INTERNATIONAL LAW & GOVERNANCE | OMIDYAR NETWORK | FLETCHER INSTITUTE FOR BUSINESS IN THE GLOBAL CONTEXT

DEFEATING DISINFORMATION

Advancing Inclusive Growth and Democracy through Global Digital Platforms

Policy Approaches to Defining and Enforcing Responsibilities for Online Platforms

Josephine Wolff

WORK IN PROGRESS

This research is published as part of the Defeating Disinformation UnConference, a joint effort by Digital Planet and the Center for International Law and Governance, at The Fletcher School at Tufts University. The Defeating Disinformation UnConference is supported by the Omidyar Network.

Policy Approaches to Defining and Enforcing Responsibilities for Online Platforms

Josephine Wolff
August 2022

In 1996, when the United States Communications Decency Act was passed, including the Section 230 provisions that protect online service providers from being held liable for most content posted on their platforms by users, the five most visited websites on the Internet were AOL, Webcrawler, Netscape, Yahoo, and Infoseek.[1] By June 2022, the list of the five most visited websites in the world consisted entirely of companies that had not even been founded in 1996, when Section 230 was adopted: Google, YouTube, Facebook, Wikipedia, and Twitter.[2] Critics of Section 230's liability protections for online platforms point out how much the online landscape has shifted in order to argue that the law was written without any understanding of the new types of platforms that would emerge on the Internet and the impacts they would have. Supporters of Section 230, by contrast, point to these new platforms as evidence of the law's success, arguing that websites like YouTube, Facebook, and Twitter could never have existed without the protections enshrined in Section 230 because their developers would otherwise have faced crushing liability risks. What's clearly true is that the dominant platforms of the modern Internet look very different from the ones that were most popular at the time that Section 230 was passed, and that these newer platforms are now giving rise to a wide range of very different policy approaches around the world. Whether the emergence of these platforms is viewed as a positive consequence of an unexpectedly prescient liability regime underpinned by Section 230, or a negative consequence of a disastrously lax approach to platform liability, there is no question that they are now subject to an increasingly complicated—and growing—set of public policies that alternately target specific types of platforms, specific types of content, and specific types of liability in the context of a variety of different, competing policy priorities.

At the heart of this proliferating set of platform policies and proposals is not just the global growth of the Internet since 1996—though that has, of course, driven many more countries to take an interest in regulating online platforms—but also the fact that the individual platforms themselves now serve many more functions than they used to, ranging from providing access to news and information to providing encrypted messaging systems to selling cloud storage. This means that the types of content hosted on each of these platforms are more varied, the groups of people using them are more diverse, and the threats they pose are much more numerous and difficult to predict and disentangle. Accordingly, online platform regulation now involves multiple different strands of policy focused on different types of responsibilities for different types of platforms—a stark contrast to the one-size-fits-all approach of Section 230 which granted liability protections with a fairly broad brush to all "interactive computer services," with the exception of some criminal and intellectual property-related claims.

This paper aims to build on the five comparative country briefs as well as the three generative

[1] A LOOK BACK IN TIME... AT THE MOST VISITED WEB DOMAINS OF 1996!, https://www.comscore.com/Insights/Blog/A-Look-Back-in-Time-at-the-Most-Visited-Web-Domains-of-1996
[2] TOP WEBSITES, https://www.semrush.com/website/top/

briefs authored for this project by breaking down and disentangling some of the different policy debates and goals surrounding online platform responsibility. Identifying these different threads in the technology policy domain allows for a more granular discussion of which, if any, of these goals might lend themselves to some greater degree of international coordination, and whether there is potential for global governance solutions around certain issue areas related to online platforms. Similarly, separating out different policy concerns is useful for considering the implications of the generative papers about banking regulations, taxation, and public health. As the authors of those briefs discuss, none of those areas offers a perfect analogy to the problems posed by online platforms, but each of them may still have something to offer by way of insight with regard to particular facets of online platform regulation.

This paper is structured as follows. The next section reviews key themes of the comparative briefs with the goal of identifying different policy goals related to online platforms along three different types of obligations: responsibilities to target particular categories of unwanted content, responsibilities for platforms that wield particularly significant influence, and responsibilities to be transparent about platform decision-making. The last section considers which of these policy goals present the greatest opportunities for international coordination and agreement and which of them actually require such coordination in order to be effectively implemented. Finally, it considers what lessons can be drawn from existing policy efforts for how to foster greater coordination around areas of common interest related to online platforms.

Goals of Online Platform Policy Making

The goals of regulating platform providers and their responsibilities have evolved and expanded considerably since the passage of Section 230. Jeff Kosseff traces the history and motivations of Section 230 to identify the primary goals that policymakers had in mind when authoring those early protections for online platforms and finds that one of the main reasons Congress passed that section of the law was to enable websites to be able to perform content moderation without facing additional liability for doing so. Kosseff explains that prior to the passage of Section 230, many U.S. courts held that websites that had reason to know about content posted on their platforms could be held liable for it. So, when the dial-up service provider CompuServe was sued because an online newsletter it distributed contained allegedly libelous material, a court dismissed the case in 1991 on the grounds that company had no way of knowing about that material. But when CompuServe's competitor Prodigy was later sued for comments made on one of its financial bulletin boards, a court found that it had opened itself up to greater liability than CompuServe by voluntarily hiring moderators to review posts on its platform and enforce user conduct rules.[3]

The Prodigy case attracted the attention of then-U.S. Representatives Chris Cox and Ron Wyden, Kosseff explains, and in 1995 they introduced a bill that would later turn into the core of the Section 230 protections, with the intention—according to the law itself—of "promot[ing] the continued development of the Internet and other interactive computer services and other interactive media" as well as "remov[ing] disincentives for the development and utilization of blocking and filtering technologies that empower parents to restrict their children's access to objectionable or

[3] JEFF KOSSEFF, THE TWENTY-SIX WORDS THAT CREATED THE INTERNET (Cornell University Press) (2019)

inappropriate online material."[4] Kosseff notes, "From the relatively sparse legislative history, it is clear that Section 230's drafters had two primary goals. First, they wanted to ensure that the nascent commercial Internet was unburdened from regulation and litigation. Second, they wanted to encourage online providers to moderate as they (and their users) saw fit."[5]

These aims—especially the goal of encouraging Internet innovation and development—were not uncommon priorities for regulators during the late 1990s and early 2000s, even outside of the United States. Indeed, the template laid out by Section 230 of enacting broad liability protections for online service providers and carving out a couple specific exceptions was echoed in Brazilian policy, which provided safe harbor for platforms covering all content besides copyright claims and non-consensual intimate content (as well as anything a court explicitly ordered them to remove), as Monteiro's brief points out.[6] Similarly, Indian platform policy has for many years also revolved around the promise of liability protections for platforms except in the case of two types of content: non-consensual intimate content and rape and child sex abuse materials.[7] However, not all countries took this model as their basis for shaping platform liability. Busch's paper describes how the European Union's E-Commerce Directive—a precursor to the Digital Services Act—took exactly the approach that led to the difference in treatment of CompuServe and Prodigy by dictating that only platform operators that were not playing any "active role" in determining which content was permitted on the platform were protected from being held liable for hosting that content. The Digital Services Act also provides platforms with liability protections conditional on the platform not knowing about the illegal content they host, but it adds a "Good Samaritan rule" that aims at the same goal as Section 230: permitting providers to conduct moderation activities and self-initiated investigations without losing their liability protections.[8]

Content-Specific Carve-Outs

These twin goals of promoting the online economy and also enabling online platforms to moderate user-generated content without facing additional legal burdens for doing so were important in shaping the platform policies of many countries—but not all. In China, as Jufang Wang's paper points out, online platforms have instead been assigned "primary responsibility" for the content that users post on their servers and are therefore required to constantly monitor and moderate that content.[9] The policy goal motivating this model is presumably control over the types of information being shared via online platforms, and while China takes a particularly strong stance in prioritizing that as their top policy aim with regard to platforms, almost all of the countries that have passed platform-related regulations have made it a priority to force platforms to police certain types of content proactively. Most of these priorities have emerged as targeted carve-outs from broad platform liability protection laws, such as the exemptions in Section 230 for criminal liability and copyright-related claims, or for copyright and non-consensual intimate content in Brazil, or for non-consensual intimate content and rape and child sex abuse materials in India.

[4] 47 U.S.C. § 230 PROTECTION FOR PRIVATE BLOCKING AND SCREENING OF OFFENSIVE MATERIAL
[5] TESTIMONY OF JEFF KOSSEFF BEFORE THE SENATE COMMITTEE ON COMMERCE, SCIENCE & TRANSPORTATION, https://www.commerce.senate.gov/services/files/444EFF87-84E3-46DB-B8DB-24DC9A424869
[6] BRIEF ON PLATFORM RESPONSIBILITY IN BRAZIL
[7] COMPARATIVE PAPER—INDIA
[8] PLATFORM RESPONSIBILITY IN THE EUROPEAN UNION
[9] PLATFORM RESPONSIBILITY WITH CHINESE CHARACTERISTICS

By carving out specific categories of content that platforms could be held liable for, countries could offer fairly broad liability protections, reducing the burden on online service providers to monitor everything, while still maintaining a few core priorities for certain types of content that were of particular concern. Unsurprisingly, these priorities varied by country—but they have also changed and evolved over time, alongside the growth and evolution of the online platforms themselves. In the 1990s, the existence of popular peer-to-peer file-sharing websites like Napster, as well as a robust music and film industry with a significant lobbying presence in Washington, DC, meant that liability for illegal sharing of copyrighted material was a particular concern in the United States, leading to the Section 230 exemption for copyrighted material. The policies established by Brazil and India, developed later on and in the context of very different political climates, prioritized liability for non-consensual intimate content. And US regulators, in revisiting the protections of Section 230 in recent years in light of concerns about online sex trafficking, misinformation, and censorship, have introduced proposals that target the liability protections for a variety of other categories of content, including inaccurate information, political speech, advertisements. Of these myriad proposals, only one significant carve out to the Section 230 protections has actually been enacted in the United States—the controversial Stop Enabling Sex Traffickers Act (SESTA) and Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA), both passed in 2018, which exempted federal and state sex trafficking laws from Section 230's liability protections.[10] SESTA/FOSTA caused many websites, including Craigslist, Reddit, and Facebook, to stop hosting online advertisements for sex work or other sexually explicit content for fear of any resulting liability, leading critics of the law to allege that rather than helping crack down on sex trafficking, it had only made sex work more dangerous and difficult.[11]

Other countries have similarly introduced specific carve-outs to their liability protections that center on specific types of content. For instance, Busch describes the provisions governing "terrorist content" in the European Union, and Monteiro describes the proposed Brazilian policy that would have made platforms legally liable for misinformation posted on their platforms.[12] India has similarly introduced a variety of content-specific liability policies, including ones for terrorism-related content and election-related content.[13] Interestingly, India's approach to regulating liability protections for platforms has hinged in some cases not just on the type of content but also the technical transmission mode and format of that content. Specifically, Indian regulators have taken aim at end-to-end encrypted messaging platforms in their 2021 Intermediary Guidelines and Digital Media Ethics Code by requiring platforms to "enable the identification of the first originator" of any message—something that is technically infeasible under end-to-end encrypted systems.[14] This intersection of platform liability policy with encryption policy highlights how intertwined the various areas of Internet policy have become in light of the many different functions that online platforms now perform. By targeting specific types of content and imposing additional obligations and liability on platforms related to those categories, regulators can

[10] H.R.1865 - ALLOW STATES AND VICTIMS TO FIGHT ONLINE SEX TRAFFICKING ACT OF 2017, https://www.congress.gov/bill/115th-congress/house-bill/1865/text; S.1693 - STOP ENABLING SEX TRAFFICKERS ACT OF 2017 , https://www.congress.gov/bill/115th-congress/senate-bill/1693/text

[11] David McCabe & Kate Conger, *Stamping Out Online Sex Trafficking May Have Pushed It Underground*, THE NEW YORK TIMES, Dec. 17, 2019, https://www.nytimes.com/2019/12/17/technology/fosta-sex-trafficking-law.html

[12] Monteiro; Busch

[13] Comparative Paper — India

[14] INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE, https://egazette.nic.in/WriteReadData/2021/225464.pdf

sometimes strike a balance between the broad goals that motivated Section 230, as well as other early platform policies, and the specific policy priorities that revolve around particularly controversial or volatile issues or types of content. However, unless these carve-outs are very narrowly targeted they can sometimes backfire with platforms choosing to over-zealously remove content they fear may open them up to greater liability, as in the case of SESTA/FOSTA and some U.S. digital copyright protections, or finding that the technical implementation of their services is incompatible with the requirements for their liability protections, as in the case of the Indian Intermediary Guidelines.

Notably, most of these carve-outs focus on types of information that regulators want to discourage platforms from leaving up by imposing additional liability on them if that content is not removed. However, carve-outs to liability protections can also work in the opposite way: requiring platforms *not* to take down certain types of content or be liable for its removal. For instance, in September 2021, Brazilian president Jair Bolsonaro temporarily banned social media companies from removing content related to the country's 2022 elections, including his own posts claiming that he would win the election unless the voting machines were compromised. The Brazilian ban forbade platforms from removing any content that did not involve nudity, drugs or violence and required them to obtain court orders before being permitted to remove any other types of content from their platforms.[15] Carve-outs of this nature reverse the motivation for policies like Section 230, aiming to disincentivize content moderation by increasing the potential liability associated with proactively moderating online posts, rather than encouraging that voluntary moderation activity by providing liability protections for it.

Platform-Specific Liability Regimes

Both of the policy goals that motivated Section 230—promoting development of the Internet economy and protecting platforms that choose to engage in voluntary content moderation—have shifted somewhat over the course of the past decade, though neither has disappeared. As the online economy has grown and the major digital platforms have become more established, regulators' fears of inhibiting the continued development of the Internet have receded somewhat, at least in countries where there is sufficient Internet penetration and infrastructure. In the United States, for instance, it is no longer plausible to argue that without broad liability protections for online platforms there will be no successful platforms, especially given how many resources the most successful of those companies now have to put towards compliance. However, there is still considerable concern in many countries with developed tech sectors that they may fall behind other countries or fail to innovate as quickly and successfully as their competitors. So the policy goals that now motivate the preservation of regulations like Section 230 and analogous laws in other countries center on the importance of innovation and protecting new and emerging technologies and companies from high barriers to entry. Facebook or Google might be able to afford to comply with onerous liability regimes, advocates for broad liability protections now argue, but newer, smaller companies won't be able to, so any attempt to impose more liability on platforms will only serve to further entrench the existing dominant platforms and hinder innovation and competition from new entrants.

[15] Jack Nicas, *Brazil's President Bans Social Networks From Removing Some Posts*, THE NEW YORK TIMES, Sep. 9, 2021, https://www.nytimes.com/2021/09/09/world/americas/bolsonaro-social-networks.html

Whether or not these concerns are warranted, they have yielded a newer style of platform regulation in recent years—one that focuses specifically on the largest platforms, carving up the ever-growing landscape of online intermediaries to pin additional responsibilities on only the largest and most influential platforms. Busch describes the DSA Level 4 due diligence obligations that apply only to "very large online platforms" (VLOPs) and "very large online search engines" (VLOSEs) that have more than 45 million average monthly users in the EU. These obligations include conducting risk assessments and independent annual audits of DSA compliance at their own expense, granting researchers and regulators access to data related to systemic risks, publishing their terms and conditions in the official languages of every country where they offer their services, making the "main parameters" used by their recommender systems "clear and accessible," and more frequent transparency reports (every six months instead of every year). The rationale for these additional obligations is explained in the Act itself which states that it "sets asymmetric due diligence obligations on different types of digital service providers depending on the nature of their services and their size" because "[c]ertain substantive obligations are limited only to very large online platforms, which due to their reach have acquired a central, systemic role in facilitating the public debate and economic transactions" and "which are prone to the highest levels of risks for the EU society and economy." By contrast, the DSA states, "[v]ery small providers are exempt from the obligations altogether."[16] Even the lowest level of DSA obligations exclude all "micro" and "small" enterprises, or all enterprises with fewer than 50 employees and annual turnover of less than 10 million Euros.[17]

In its Intermediary Guidelines, India similarly defines a category of "significant social media intermediaries" (SSMIs) that have more than 5,000,000 users in India. These SSMIs, like VLOSEs and VLOPs have additional obligations under the Guidelines, including being required to appoint a Chief Compliance Officer as well as a contact for 24/7 coordination with law enforcement officials, to publish monthly compliance reports, and to identify the originator of all messages sent via its platform.[18] In contrast, the United States and China have largely avoided imposing different obligations to larger or smaller platforms, choosing instead to regulate all online service providers consistently, perhaps in part because they are home to some of the largest platform companies who most strenuously object to being the subject of additional responsibilities. The goals of policy approaches tailored to the size of platforms are generally to impose requirements and obligations that are proportional to the resources of individual platforms, thereby placing the greatest burden of compliance on the companies that are both best able to shoulder that burden and also most likely to have the greatest impact on the largest number of users. In other words, like Section 230 and its successors, these policies are partly motivated by a desire to encourage innovation—and not discourage the creation of new online start-ups by requiring them to undertake onerous compliance regimes—and also partly motivated by a desire to encourage as much proactive content moderation as possible from the companies that can afford it.

Transparency Obligations

[16] DIGITAL SERVICES ACT, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0825&from=IT

[17] *COMMISSION RECOMMENDATION of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises*, OFFICIAL JOURNAL OF THE EUROPEAN UNION (2003), https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:124:0036:0041:en:PDF

[18] Ministry of Electronics and Information Technology

A final category of policy approaches to platform responsibility center on transparency obligations of the online platforms. These include the DSA provisions requiring platforms to disclose "policies, procedures, measures and tools used for the purpose of content moderation," the transparency reports detailing how many complaints SSMIs received each month and how many pieces of content they removed that are mandated by India's Intermediary Guidelines, or the proposed Brazilian fake news bill that would require platforms to disclose sponsored online content. In the United States, many of the largest platform providers publish transparency reports about the quantity of content they take down and who they receive takedown requests from, however these reporting mechanisms are largely voluntary. In many cases, this reporting activity dates back to the transparency reporting that several tech platforms began doing in the wake of the Edward Snowden leaks to try to reestablish trust with their users by providing them with as much information as they were legally allowed to about when they turned over user data or otherwise complied with government orders. Legislative proposals that would require more transparency around online advertisements have not met with much success in the United States, though the Federal Trade Commission does require the labeling of ads and sponsored content on many platforms for consumer protection purposes.[19] In China, platform transparency reports are less common, according to Wang, but companies like TikTok that are trying to maintain a global user base publish them in order to align with expectations for Western platforms.[20]

The goals of these transparency requirements range from informing users about the nature of the content they view on these platforms—in particular whether someone is paying to show it to them and, if so, whom—to helping users understand the rules and mechanisms for content moderation. The emphasis of what, precisely, must be made transparent and to whom is dictated by the primary concerns of the regulators in question. For instance, in the DSA, EU regulators prioritize transparency around recommender systems and automated content moderation because of their concerns that automation will render these systems too difficult for users to comprehend. In India, the emphasis is on revealing how much content is removed and why, due to concerns about over-zealous or insufficient moderation (or both). In Brazil and the United States, meanwhile, transparency proposals for platforms often intersect with concerns about misinformation and aim to mitigate the impact of information on platforms by forcing disclosures about when someone has paid, or been paid, to post that content so that users have more context in which to evaluate the information they are exposed to.

Opportunities for International Cooperation & Learning from Other Domains

The motivations for online platform-related policy making span a wide range of policy goals from encouraging innovation to supporting content moderation of unwanted content, preventing the removal of desired online content, and better informing users about when content is removed and why, as well as the nature of the content they are viewing. Reconciling all of these policy goals is a challenge even within the borders of one country, much less at the international level, because some of these goals are directly at odds with each other, like encouraging voluntary moderation to reduce unwanted content and insisting platforms can't remove most forms of content without a

[19] FTC RELEASES ADVERTISING DISCLOSURES GUIDANCE FOR ONLINE INFLUENCERS, https://www.ftc.gov/news-events/news/press-releases/2019/11/ftc-releases-advertising-disclosures-guidance-online-influencers

[20] Wang

court order to reduce unwanted censorship of political speech by platforms. For the major platforms operating across multiple countries, there might be value in greater international harmonization of the rules governing some of these areas, but to a large extent the platforms themselves have been a primary instrument of such harmonization by implementing some of the same sets of community guidelines for acceptable content or disinformation labeling practices across all of the jurisdictions in which they operate. Indeed, for platforms like Facebook, when content moderators decide whether or not to remove flagged posts, they first check whether a post violates the company-wide Community Standards for content and only refer to the laws governing the jurisdiction where the user appears to be located if the content passes that first test. In other words, while platforms still expend time and resources complying with diverse, uncoordinated national laws, they have attempted in many cases to focus more of their resources on implementing a single, company-wide set of rules for content moderation that takes precedence over different national laws.

International Cooperation

Given how culturally- and context-specific rules and ideas about illegal speech, misinformation, and censorship are, it is difficult to imagine how countries would reach any clear consensus on the right way to regulate these issues or why it would be beneficial for them to do so when global platforms are, on the whole, able to tailor moderation practices to different national versions of their platforms. For instance, many large platforms, including Facebook and Google, maintain different country-specific versions of their websites for users that can be tailored to the domestic laws of that country. So when a user visits one of these websites, they will be automatically directed to the version that corresponds to and adheres to the local laws of the country where their connection and device appears to be located, based on information like their IP address, their self-identified home address, and other device and connection information. This data is collected and analyzed not just to enable platforms to comply with domestic laws, but also so they can target users with ads based on their location. In other words, the analysis of where a user is does not add a significant burden for the platform beyond what they would do anyway to provide location-based ad targeting.

This automatic redirection to different versions of a service can be reasonably effective—though the determinations made about where a user is geographically using location information are not always correct—but it does not actually prevent people in one country from accessing another country's version of the same platform. Therefore, one downside of this approach is that it leaves unwanted content available on other countries' instances of those same platforms—meaning that it is often still accessible to people even in the country where it has ostensibly been removed if they are willing to put in some effort to locate it. On the other hand, this can also be viewed as an advantage of the geofencing approach to content moderation: one country's rules about what platforms may or may not host online need not dictate the content available to other people in other places.

Indeed, just as U.S.-based platforms and tech companies like Apple are sometimes criticized for acceding to the demands of other countries, especially China, about what content and software to offer users, Wang's case study of TikTok makes clear that the reverse is also true: China-based platforms hoping to attract a global user base feel pressure to comply with standards and

expectations of users and governments in other countries. This represents, in a sense, a certain sort of international standardization: multi-national platforms segment their online presence by country and apply the national laws of each country to the specific instance of their platform aimed at that audience, directing users to the appropriate version of the site based on the location of their IP address while still allowing them to deliberately navigate to other versions. Wang's example of TikTok publishing transparency reports to align with Western companies indicates that this approach extends even beyond the companies' online presences to their other activities and public facing releases. While enforcing different rules and policy measures for different national versions of a platform is certainly not the same as reaching international consensus on these measures, it is no small thing that so many countries have accepted this approach as a satisfactory implementation of their domestic laws (there are of course notable exceptions of countries that instead choose to block entire platforms rather than allow tailored, local versions). This widespread acceptance of a highly imperfect system of implementation and enforcement is, itself, an impressive feat of international coordination.

Unsurprisingly, the areas where there have been the most successful efforts towards international cooperation on specific policy goals and measures are those where there is widespread agreement about what types of information are illegal or unwanted. For instance, platforms and law enforcement officials around the world coordinate on efforts to remove child sexual abuse materials, working from a shared database of image and video hashes. Similarly, there has been considerable international coordination around the policing and removal of copyrighted content on online platforms. This coordination has been possible not just because there are well aligned laws on these issues in many countries, but also because the platforms themselves have not had to worry about alienating large swaths of their user bases by removing these classes of content. Political content and disinformation, by contrast, has been trickier to coordinate internationally and also—not unrelatedly—harder to moderate in ways that don't infuriate large groups of users.

Interestingly, several large platforms have taken more proactive steps to moderate disinformation around public health topics, including the Covid-19 pandemic, than they have to remove disinformation concerning political candidates or topics. In large part this seems to stem from the perception within the platforms that there is clear, agreed-upon expert consensus on issues relating to topics like Covid-19 that they can use to distinguish disinformation, but no way to assemble an unbiased expert panel to make similar distinctions around political information. This points to another possible opportunity for international cooperation around non-political disinformation campaigns that might allow for more shared strategies and standards for what types of content should be removed from online platforms through consultation with international groups of experts.

Lessons from financial regulation

The generative papers for this project propose possible parallels between platform regulation and global governance schemes for other domains. Having disentangled some of the different policy goals surrounding platform responsibility-focused policies, it is perhaps useful to consider which of these parallels are most relevant to different aims. For instance, Lupo-Pasini proposes in his discussion of international financial regulation that one of the lessons tech regulators might draw from the Basel Accord is that "regulatory standards became increasingly stringent and more

detailed" over time. In that context, he suggests, it might make sense for policymakers "to work on a platform responsibility framework that is flexible enough to entice reluctant regulators to join in while progressively working to strengthen it once regulators and the industry have become used to it."[21]

This is an approach that seems best aligned with the areas where there is broad consensus around a rough policy goal but no clear agreement about the correct implementation. For instance, the policy efforts driving more transparency around how platforms moderate content and who pays for content on those platforms might be susceptible to a framework of this sort, whereby countries agree to promote some degree of reporting or transparency for platforms (as many of them already do) and then gradually arrive at a more standardized, specific set of transparency requirements over time. This approach also gives companies and regulators more time to figure out what types of reporting and transparency related to recommender systems and automated content moderation are helpful and understandable for a general audience.

It's less clear whether the lessons drawn from financial regulation about who will advocate for international policy efforts will apply to online platforms, however. Lupo-Pasini points out that in the case of the Basel Accord, "there was pressure from the broader financial industry to implement the standards. Hedge funds, banks, investment funds, rating agencies, and other market players shared the common belief that capital adequacy rules were necessary to guarantee financial stability. Therefore, regulators faced much market pressure to adopt the standards as part of their prudential toolbox if they wanted to attract foreign capital."[22] Many of the policy goals related to online platform responsibilities do not have a strong lobby in the private sector—indeed, the one that did (protecting copyrighted material and keeping infringing content off online platforms) was carved off into its own set of policies in the late 1990s and has been much more rigorously enforced than most other policy efforts in this space. The policy goals of promoting moderation, requiring platforms to remove unwanted user content, preventing platforms' from removing user content, placing additional due diligence burdens on the largest platforms, and imposing transparency requirements on platforms all have no strong support from the online platforms or any other powerful private sector coalition. Potentially, industry stakeholders might seek greater international harmonization of rules in these areas in order to reduce the need for them to customize moderation practices and national versions of their services to different jurisdictions, but so far it is not clear that this is a priority for the big platforms. Many of them have to employ separate teams to build and moderate individual national versions of their services anyway due to the differences in language, so more harmonization of rules might not actually provide such substantial relief from the burdens of trying to provide a service to people in countries all over the world.

Moreover, while at least some of these broad policy goals are held by many governments, unlike in the lessons drawn from international tax efforts, there are no clear financial incentives for governments to push these policy priorities forward, much less for them to work together to do so. Policy measures like the EU's DSA, which allows regulators to issue fines of up to 6 percent of global annual turnover for companies that violate the Act, could potentially generate some income for the governments imposing them. However, these types of penalties in EU technology regulations are often seen as a means for the EU to extract money from the primarily US-based

[21] PLATFORM RESPONSIBILITY REFORMS: LESSONS FROM INTERNATIONAL FINANCIAL REGULATION
[22] Lupo-Pasini.

major tech companies and can thereby actually serve to discourage cooperation between the United States and Europe. This suggests that at least in some cases, reaching greater international cooperation on platform-related policies may necessitate governments giving up some of the financial incentives that might otherwise help motivate the passage of such policies in a strictly domestic or regional context.