THE ILLUSION OF INCLUSION: THE FALSE PROMISE OF THE NEW GOVERNANCE PROJECT FOR CONTENT MODERATION

*Brenda Dvoskin**

Because private companies now control the most prominent communication platforms, the most pressing question in the field of content moderation is how to ensure that the governance of public discourse responds to public values. The prevailing approach, given that the state cannot regulate speech directly, is that state regulation can be substituted with audited self-regulation, broad stakeholder participation, and negotiated rulemaking. In this model, which this Article refers to as the "new governance model for content moderation," companies include advocates as representatives of the public in their processes to govern online speech. Ideally, they negotiate policy goals and share responsibility for achieving them. The end goal is to have a process in which public values are given effect. This Article argues, however, that this governance model is unsound in both theory and practice. In the field of content moderation, the ambition of constructing public values through a collaborative process between companies and stakeholders is conceptually incoherent: those interests that cannot elicit the cooperation from corporate actors and are not consistent with the values of participating advocates are excluded by design. In practice, no present or past demonstrations have shown that the inclusion of advocates in speech governance and the agreements they reach with companies have epistemic credibility to construct the public interest. Though those flaws might seem unsurprising, scholars and activists double down on independence, diversity, and expertise as design strategies that can result in

* Associate Professor of Law, Washington University in St. Louis. Thanks to Aziza Ahmed, Jack Balkin, Samantha Barbas, Yochai Benkler, Elettra Bietti, Robyn Caplan, Guy-Uriel Charles, Julie Cohen, Evelyn Douek, Noah Feldman, Janet Halley, Luke Herrine, Thomas Kadri, Amy Kapczynski, Paddy Leerssen, Orly Lobel, Paul Ohm, Robert Post, Morgan Ricks, Alan Rozensthein, Joshua Sellers, Ganesh Sitaraman, Scott Skinner-Thompson, Fred Smith, and Ari Waldman. I also thank the participants at the Privacy Law Scholars Conference, Platform Research Network, Freedom of Expression Scholars Conference at Yale Law School, LPE Conference at Harvard Law School, Networks, Platforms & Utilities (NPU) workshop, and workshops at Cornell Tech, the Institute for Information Law at Amsterdam University, and Torcuato Di Tella University. Special thanks to Ryan Miller and the amazing team at the *Fordham Law Review*.

self-regulatory bodies that could adequately set policy goals. This Article advocates for pluralism as a framework that more effectively achieves the participatory goals of new governance. It argues that the state has a central role to play in creating a plural and contested public sphere. A robust legal system can complement self-regulation and push it structurally in the direction of public values.

INTRODUCTION

The most pressing challenge in the field of content moderation today is devising a model ensuring that the governance of public discourse advances public values even though a handful of private companies control the most prominent communication platforms. Historically, the march of technology—from the printing press to film, radio, and television—has disrupted settled social structures for the production and distribution of speech.[1] Most recently, the internet revolutionized communications, and soon after, a new configuration of incumbents emerged: a few corporate actors now control the flow of digital information.[2]

1. Yochai Benkler, *Power and Productivity: Institutions, Ideology, and Technology in Political Economy*, *in* A POLITICAL ECONOMY OF JUSTICE 27, 37 (Danielle Allen, Yochai Benkler, Leah Downey, Rebecca Henderson & Josh Simons eds., 2022); HAROLD A. INNIS, THE BIAS OF COMMUNICATION (1951) (examining the relationship between the development of communication media and society); ELIZABETH L. EISENSTEIN, THE PRINTING PRESS AS AN AGENT OF CHANGE (1979) (examining the social impact of the printing press); JAMES R. BENIGER, THE CONTROL REVOLUTION: TECHNOLOGICAL AND ECONOMIC ORIGINS OF THE INFORMATION SOCIETY (1986) (examining the impact of new technology on the social order between 1840 and 1920); ERIK BARNOUW, A TOWER IN BABEL: A HISTORY OF BROADCASTING IN THE UNITED STATES TO 1933 (1966) (examining the development of radio and television and their impact in political and social life); YOCHAI BENKLER, ROBERT FARIS & HAL ROBERTS, NETWORK PROPAGANDA: MANIPULATION, DISINFORMATION, AND RADICALIZATION IN AMERICAN POLITICS (2018) (examining the impact of cable television).

2. *See* LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE, VERSION 2.0, at 138 (2006) (anticipating the recentralization of the web). *See generally* Yochai Benkler, *Degrees of Freedom, Dimensions of Power*, 145 DÆDALUS 18 (2016) (mapping the new points of centralization in the internet infrastructure); Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296 (2014) (analyzing the power of companies over digital speech).

A major obstacle to addressing this excessive accumulation of corporate power is that the law is significantly barred from directly regulating speech.[3] Recently, in *Moody v. NetChoice*,[4] Justice Kagan writing for the majority of the U.S. Supreme Court underscored that mandates to host or exclude content from social media platforms often unconstitutionally interfere with companies' editorial discretion.[5] Importantly, that does not mean that the First Amendment precludes the law from meaningfully intervening in the design of the public sphere—as Justice Alito's concurrence and Part IV of this Article argue.[6] Yet, the highly deregulatory view of the First Amendment that Justice Kagan embraces in her opinion partially explains why the dominant proposals for platform content governance deprioritize the use of law and instead tout audited self-regulation, broad stakeholder participation, and negotiated rulemaking as the best path forward.[7]

I refer to these proposals as the "new governance project for content moderation" because their premises take inspiration from the new governance school of thought, which at the beginning of this century sought to reimagine governance as a series of networked "negotiated relationships."[8] This model promises to infuse social media governance with public values by asking corporations and advocates—civil society organizations, experts, and others outside the state and corporations who have an interest in influencing how speech is produced and distributed—to collaborate in defining and achieving policy goals.[9] This Article examines the new

3. *See, e.g.*, Noah Feldman, *Facebook Supreme Court: A Governance Solution*, *in* GLOBAL FEEDBACK AND INPUT ON THE FACEBOOK OVERSIGHT BOARD FOR CONTENT DECISIONS APPENDIX 101, 102 (2018), https://about.fb.com/wp-content/uploads/2019/06/oversight-board-consultation-report-appendix.pdf [https://perma.cc/3EHB-QBVE] ("The U.S. government can't play that role for U.S.-based companies, because those companies are themselves entitled to the freedom of speech. Legally, that means the government can't impose limits on what Facebook or Google choose to allow on their platforms. What they choose to allow is itself an expression of their own First Amendment rights."); Evelyn Douek, *Content Moderation as Systems Thinking*, 136 HARV. L. REV. 526, 603 (2022).

4. 144 S. Ct. 2383 (2024).

5. *Id.* at 2405–06.

6. *See id.* at 2431–33 (Alito, J., concurring in the judgment); *infra* Part IV.

7. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1666 (2018); Evelyn Douek, *Verified Accountability: Self-Regulation of Content Moderation as an Answer to the Special Problems of Speech Regulation* 1–2 (Hoover Inst., Aegis Series Paper No. 1903, 2019); Hannah Bloch-Wehba, *Global Platform Governance: Private Power in the Shadow of the State*, 72 SMU L. REV. 27, 71 (2019); Jillian C. York & Ethan Zuckerman, *Moderating the Public Sphere*, *in* HUMAN RIGHTS IN THE AGE OF PLATFORMS 137, 157 (Rikke Frank Jørgensen ed., 2019); Rory Van Loo, *Federal Rules of Platform Procedure*, 88 U. CHI. L. REV. 829, 832 (2021); Douek, *supra* note 3, at 593; Martha Minow & Newton Minow, *Social Media Companies Should Pursue Serious Self-Supervision—Soon: Response to Professors Douek and Kadri*, 136 HARV. L. REV. F. 428, 432 (2023); Gilad Abiri & Sebastián Guidi, *The Platform Federation*, 26 YALE J.L. & TECH. 240, 245–49 (2024).

8. Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543, 571 (2000); Kate Klonick, *Of Systems Thinking and Straw Men*, 136 HARV. L. REV. F. 339, 350 (2023) (identifying new governance as a source of inspiration).

9. Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 MINN. L. REV. 342, 344 (2004).

governance model as it has been imported and applied to the field of content moderation.[10]

Some of the projects designed within this frame have been gaining great traction. Partially implementing this vision, Mark Zuckerberg announced the Meta Oversight Board (the "Board") in 2018, a body currently constituted by twenty-three experts in law, journalism, and human rights that reviews Meta's content moderation decisions.[11] Soon after, Twitter revamped its Trust and Safety Council—now dismantled—and TikTok, Twitch, and Spotify launched their own advisory boards constituted by prominent users, members of nonprofits, and experts in diverse fields.[12] All major tech companies regularly engage advocates as a standard element in their process of writing content moderation rules.[13] Other multistakeholder initiatives have taken on various tasks. The Global Internet Forum to Counter Terrorism brings together representatives of companies, civil society, and academia to reach agreements on how to govern terrorist content.[14] The Global Network Initiative develops frameworks to guide the relationships

10. I thank Daphna Renan and Molly Brady for a conversation on this point.

11. Mark Zuckerberg, *A Blueprint for Content Governance and Enforcement*, FACEBOOK (Nov. 15, 2018), https://www.facebook.com/notes/mark-zuckerberg/a-blueprint-for-content-governance-and-enforcement/10156443129621634 [https://perma.cc/BK23-H2UK]; *see also* Kate Klonick, *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*, 129 YALE L.J. 2418 (2020) (following the creation of the Oversight Board).

12. *Strengthening our Trust and Safety Council*, X BLOG (Dec. 13, 2019), https://blog.x.com/en_us/topics/company/2019/strengthening-our-trust-and-safety-council.html [https://perma.cc/64BU-6VU8]; Vanessa Pappas, *Introducing the TikTok Content Advisory Council*, TIKTOK (Mar. 18, 2020), https://newsroom.tiktok.com/en-us/introducing-the-tiktok-content-advisory-council [https://perma.cc/8KFB-5RZG]; *Introducing the Twitch Safety Advisory Council*, TWITCH BLOG (May 14, 2020), https://blog.twitch.tv/en/2020/05/14/introducing-the-twitch-safety-advisorycouncil/ [https://perma.cc/W6UB-FLBX]; *Introducing the Spotify Safety Advisory Council*, SPOTIFY: FOR THE RECORD (June 13, 2022), https://newsroom.spotify.com/2022-06-13/introducing-the-spotify-safety-advisory-council/ [https://perma.cc/7JZZ-MYE8]. Elon Musk dissolved Twitter's Trust and Safety Advisory Council shortly after buying Twitter. Matt O'Brien & Barbara Ortutay, *Musk's Twitter Disbands its Trust and Safety Advisory Group*, ASSOCIATED PRESS (Dec. 13, 2022), https://apnews.com/article/elon-musk-twitter-inc-technology-business-a9b795e8050de12319b82b5dd7118cd7 [https://perma.cc/3RZL-BXBK].

13. Robyn Caplan, *Networked Platform Governance: The Construction of the Democratic Platform*, 17 INT'L J. COMMC'N 3451 (2023) (theorizing the network of platforms and nonstate actors that participate in platform governance); Naomi Appelman & Paddy Leerssen, *On "Trusted" Flaggers*, 24 YALE J.L. & TECH 452 (2022) (examining certain partnerships between companies and civil society); Matthias C. Kettemann & Wolfgang Schulz, *Setting Rules for 2.7 Billion: A (First) Look into Facebook's Norm-Making System: Results of a Pilot Study* 23 (Hans-Bredow-Inst. Working Paper, No. 1, 2020), https://www.ssoar.info/ssoar/handle/document/71724 [https://perma.cc/T3K2-TFJM] (describing Meta's stakeholder engagement process); Peter Stern, Sarah Shirazyan & Abby Fanlo, *How Can Platform Engagement with Academics and Civil Society Representatives Inform the Development of Content Policies?: A Look at Meta's COVID-19 Misinformation Policies*, J. ONLINE TR. & SAFETY, Sept. 2022, at 1 (same).

14. *About*, GLOB. INTERNET F. TO COUNTER TERRORISM, https://gifct.org/about/ [https://perma.cc/3HTM-U7VY].

between companies and governments.[15] Stakeholder participation might soon become legally mandatory. The European Union Digital Services Act[16] (DSA) mandates the participation of civil society organizations in several forms.[17] Most proposals for legal reform emphasize transparency regimes and audit requirements.[18] These proposals expect advocates to play the fundamental role of ensuring that publicity translates into public oversight.[19]

These initiatives frame their work as advancing the interests of the public. As salient contemporary examples, Meta describes its stakeholder engagement process as being inclusive of all viewpoints, and the Meta Oversight Board claims to derive its decisions from international human rights law (IHRL)—a legal framework that evokes a global consensus.[20] This discourse aligns with the companies' interest in showing lawmakers that firms can pursue public goals without the need for state intervention. It also aligns with the core project of the new governance paradigm of designing a collaborative praxis with broad stakeholder participation that progressively evolves toward a consensus that benefits everyone.[21]

This Article argues that the new governance project for content moderation is unsound in both theory and practice. The promise of bringing all viewpoints together to negotiate outcomes that satisfy everyone's interests is incoherent.[22] Indeed, lifting the veil to look at which advocates have gained inclusion in the governance elite shows that their claim of being representative of the public interest conceals the exercise of power by narrow groups. Thus, even though the project presents itself as substance agnostic, it has substantive implications.[23] In some cases, these implications are apparent. For example, although companies publicly claim to protect users' safety, their definition of safety has been informed by the important concerns

15. *About GNI*, GLOB. NETWORK INITIATIVE, https://globalnetworkinitiative.org/about/ [https://perma.cc/9CDP-2QSF].

16. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 Oct. 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act), 2022 O.J. (L 277).

17. Martin Husovec, *Will the DSA Work?*, VERFASSUNGSBLOG (Nov. 9, 2022), https://verfassungsblog.de/dsa-money-effort/ [https://perma.cc/WWS9-WQG9] (describing the role of civil society in the Digital Services Act framework).

18. *See, e.g.*, Algorithmic Justice and Online Platform Transparency Act, S. 1896, 117th Cong. (2021).

19. *See, e.g.*, ARCHON FUNG, MARY GRAHAM & DAVID WEIL, FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY 43 (2007) (The authors analyze the potential effects of transparency mandates generally and argue that "[b]usinesses may be forced to establish new systems of monitoring, measuring, review, and reporting . . . . [D]isclosures may change their practices in response to new knowledge as well as to public pressure.").

20. Brenda Dvoskin, *Expert Governance of Online Speech*, 64 HARV. INT'L. L.J. 85, 87–88 (2023).

21. *See* Lobel, *supra* note 9, at 405–06.

22. CHANTAL MOUFFE, AGONISTICS: THINKING THE WORLD POLITICALLY 92 (2013); Nancy Fraser, *Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy*, 1 SOC. TEXT, no. 25/26, 1990, at 59.

23. *See* JULIE E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM 187 (2019) (discussing the risky distributive effects of the new governance paradigm).

of feminists fighting for privacy protections but has excluded the demands of feminists campaigning for sex workers' safety.[24] More generally, the new governance model as it relates to content moderation amplifies the power only of advocates whose projects are not only congruent with the pursuit of profit but even benefit from the centralized speech control that a concentrated market offers. Successful advocates might represent urgent, important, and valid causes and might be empowered to make positive changes. However, whether advocacy goals are compatible with the core of the industry business model is a poor criterion for defining speech governance objectives.

This account of stakeholder participation is grounded in observations of the platform-based public sphere and three historical case studies. The first case study focuses on the Edison Trust, the first association of film producers in the United States; the second on the Hollywood studios; and the third on the National Broadcasting Company (NBC) in radio.[25] In each case, the industry created an institution to include advocates in the governance of public discourse.[26] From contemporary initiatives, Meta offers the most developed and publicly reported examples. Although each story has unique features that distinguish it, their distributive effects and impact on public discourse are instructive.

These historical case studies make three contributions. First, these past experiences provide further evidence of the inherent limitations of these projects.[27] When contemporary initiatives are observed in isolation, it can be tempting to attribute their shortcomings, for example, to the imperfections of the institutional design of the Meta Oversight Board or its current composition. Yet, without exception, the inclusion of advocates in speech governance has bent governance toward an elite-defined range of relevant viewpoints. Second, unveiling the distributional effects of this governance model underscores the practical importance of this Article's conceptual claim and motivates the search for an alternative that more closely approximates the ideal of inclusion or participatory parity. Indeed, the exclusionary effects of collaborative governance could be merely hypothetical or irrelevant. In practice, this paradigm has worked to sideline voices contesting hegemonic social norms from the margins and worthy of inclusion in the public sphere.

---

24. Rosalie Gillett, Zahra Stardust & Jean Burgess, *Safety for Whom?: Investigating How Platforms Frame and Perform Safety and Harm Interventions*, 8 SOC. MEDIA + SOC'Y 1, 8–9 (2022).

25. PAUL STARR, THE CREATION OF THE MEDIA: POLITICAL ORIGINS OF MODERN COMMUNICATIONS 307 (2004) (discussing the creation of the Edison Trust and its evolution); GREGORY D. BLACK, HOLLYWOOD CENSORED: MORALITY CODES, CATHOLICS, AND THE MOVIES (1994) (examining state and private regulation of Hollywood movies); CHRISTOPHER H. STERLING & JOHN MICHAEL KITTROSS, STAY TUNED: A HISTORY OF AMERICAN BROADCASTING (3d ed. 2002) (presenting a historical account of American broadcasting).

26. CHARLES MATTHEW FELDMAN, THE NATIONAL BOARD OF CENSORSHIP (REVIEW) OF MOTION PICTURES, 1909–1922, at 20 (1977); RICHARD S. RANDALL, CENSORSHIP OF THE MOVIES: THE SOCIAL AND POLITICAL CONTROL OF A MASS MEDIUM 198 (1968); LOUISE M. BENJAMIN, NBC ADVISORY COUNCIL AND RADIO PROGRAMMING, 1926–1945, at 11 (2009).

27. Yochai Benkler & Talha Syed, *Reconstructing Class Analysis*, J. L. & POL. ECON. (forthcoming) (manuscript at 17) (using history as a mode of validation of conceptual claims to the degree that they can explain historically observed patterns).

Finally, each case study points in the direction of a potential legal intervention that could have effectively disrupted the dominant order of the public sphere. These underexplored legal possibilities inspire the positive claims of this Article.

The insight that companies partner selectively with relatively friendly advocates might be unsurprising. Indeed, it is precisely what one would expect. Yet, contemporary proposals insist on independence, diversity, and expertise as plausible tools for building self-regulatory bodies that have epistemic credibility to construct the public interest.[28] In response to the shortcomings of initiatives like the Meta Oversight Board, scholars and activists suggest that these bodies should have stronger enforcement capabilities, that the ambit of their authority should be broadened, or that their decisions ought to become subject to stricter scholarly and public scrutiny.[29] This Article argues that this enterprise is doomed from the outset. No better institutional design can make new governance a credible way of setting public objectives for content moderation. It shows that even the most ambitiously independent, diverse, and inclusive initiatives have exclusionary effects. The scope and terms of the negotiation setting determine which interests can be channeled through this process and which are left out by design.[30] Even among those included in self-regulatory initiatives, aspiring to find speech rules that serve as common ground among all viewpoints is an illusion.

In search of an alternative normative vision, this Article explores pluralism as a paradigm that might more closely approximate the ideal of inclusion.[31] Pluralism gives excluded groups the opportunity to participate in speech production without the need to elicit cooperation from powerful actors. Law has a central role to play in creating a plural public sphere. Even though the new governance paradigm frames self-regulation as an alternative to law,[32] the two systems are better understood as complementary. Indeed, law structures the infrastructure for the production and distribution of public discourse. Among other possibilities, it can reintroduce diversity in speech

28. *See, e.g.*, PAUL GOWDER, THE NETWORKED LEVIATHAN: FOR DEMOCRATIC PLATFORMS 167 (2023); Barrie Sander, *Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation*, 43 FORDHAM INT'L L.J. 939, 991 (2020); Klonick, *supra* note 11, at 2418.

29. Evelyn Douek, *The Meta Oversight Board and the Empty Promise of Legitimacy*, 37 HARV. J.L. TECH. 373, 444 (2024); Jessica J. González & Carmen Scrutato, *Everyone on Facebook's Oversight Board Should Resign*, WIRED (Apr. 17, 2021), https://www.wired.com/story/opinion-everyone-on-facebooks-oversight-board-should-resign/ [https://perma.cc/4E82-UZAL] ("[T]he board can also make policy recommendations as part of its decisions, but these are merely suggestions that Facebook can take or leave."); Flynn Coleman, Brandie Nonnecke & Elizabeth M. Renieris, *The Promise and Pitfalls of the Facebook Oversight Board: A Human Rights Perspective* (Carr Ctr. Discussion Paper Series, May 6, 2021), https://carrcenter.hks.harvard.edu/publications/promise-and-pitfalls-facebook-oversight-board-%E2%80%93-human-rights-perspective [https://perma.cc/GLX3-MBE5].

30. *See* Amy J. Cohen, *Negotiation, Meet New Governance: Interests, Skills, and Selves*, 33 L. & SOC. INQUIRY 503, 545–48 (2008).

31. Fraser, *supra* note 22, at 68.

32. *See, e.g.*, Minow & Minow, *supra* note 7, at 432.

governance, for example by encouraging the creation of a competitive market for recommendation algorithms on social media platforms.[33] In this world, platforms would host content generated by users, while people could choose from multiple recommendation algorithms to organize that content. Each option could curate online speech advancing different values and serving diverse audiences.

Although this vision presents numerous new challenges, it addresses two main deficits of the new governance project for content moderation. First, it acknowledges that companies will partner with advocates selectively. Thus, it does not require that we look at hopeless institutional design strategies (such as independent bodies, diverse composition, and expert knowledge) to make stakeholder participation a good process to construct public goals. Instead, this model situates partnerships between corporations and advocates in an institutional context that *expects* them to defend their own interests. Here, participating actors do not work toward reconciliation but engage in confrontation as political actors. Second, it offers excluded viewpoints the opportunity to participate directly in the governance of online speech—for example, by controlling one among many recommendation systems on a social media platform. Realizing this vision will require addressing many implementation questions—some of which this Article discusses. Still, it is a more promising approximation to the public sphere as a contested space where people come together to share, encounter, and confront information, ideas, arguments, and opinions on political, cultural, and other matters.[34]

The Article proceeds as follows. Part I introduces the new governance project for content moderation. Part II shows that its fundamental promises are conceptually incoherent. Part III examines three historical case studies that provide further evidence for this conclusion. Part IV looks at legal interventions that could bring together the best aspects of self-regulation and robust state support for a more plural, egalitarian, and contested public sphere.

33. Mike Masnick, *Protocols, Not Platforms: A Technological Approach to Free Speech*, KNIGHT FIRST AMEND. INST. (Aug. 21, 2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech [https://perma.cc/EWJ4-HS77]; Daphne Keller, *The Future of Platform Power: Making Middleware Work*, 32 J. DEMOCRACY, no. 3, 2021, at 168; Robert Bodle, *Regimes of Sharing: Open APIs, Interoperability, and Facebook*, 14 INFO., COMMC'N & SOC'Y, no. 3, 2011, at 320; FRANCIS FUKUYAMA, BARAK RICHMAN, ASHISH GOEL, ROBERTA R. KATZ, A. DOUGLAS MELAMED & MARIETJE SCHAAKE, MIDDLEWARE FOR DOMINANT DIGITAL PLATFORMS: A TECHNOLOGICAL SOLUTION TO A THREAT TO DEMOCRACY 3 (2021).

34. Catherine R. Squires, *Rethinking the Black Public Sphere: An Alternative Vocabulary for Multiple Public Spheres*, 12 COMMC'N THEORY 446, 448 (2002).

I. THE NEW GOVERNANCE PROJECT FOR CONTENT MODERATION

A handful of corporate actors wield excessive power over the circulation of political, cultural, and other forms of expressions.[35] Numerous problems flow from this configuration for the distribution of speech: speech governance has democratic deficits,[36] harmful speech spreads on social media platforms,[37] companies often exclude valuable speech from their platforms[38] and misrepresent their policies,[39] and ownership concentration puts competition and even democracy at risk.[40]

Given the constitutional limits on direct regulation of content moderation through the law, the dominant response to these problems has consisted of a bundle of proposals that include audited self-regulation, disclosure regimes, and negotiated rulemaking as the pillars to reform social media.[41] From this perspective, companies should operate their businesses in a way that advances public values, include advocates in their internal governance structures to define those public values, and be transparent about how they implement those values in their business operations. In their strongest formulation, these proposals follow the premises of the new governance school of thought, which at the beginning of the century emphasized broad stakeholder participation in a negotiation process to define and achieve public goals with light-touch state support.

This part introduces this paradigm, its underlying conceptualization of how law and speech governance relate to each other, and its promises for the future of content moderation.

---

35. *See* ARI EZRA WALDMAN, INDUSTRY UNBOUND: THE INSIDE STORY OF PRIVACY, DATA, AND CORPORATE POWER 233 (2021) (describing corporate domination and excessive accumulation of power in the related realm of privacy).

36. DAVID KAYE, SPEECH POLICE: THE GLOBAL STRUGGLE TO GOVERN THE INTERNET 18 (2019).

37. *See, e.g.*, Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. REV 61, 63–64 (2009) (discussing online harassment).

38. *See, e.g.*, Abdul Rahman Al Jaloud, Hadi Al Khatib, Jeff Deutch, Dia Kayyali & Jillian C. York, *Caught in the Net: The Impact of "Extremist" Speech Regulations on Human Rights Content*, ELEC. FRONTIER FOUND. (May 30, 2019), https://www.eff.org/wp/caught-net-impact-extremist-speech-regulations-human-rights-content [https://perma.cc/VQ5L-R53E]; Ari Ezra Waldman, *Disorderly Content*, 97 WASH. L. REV. 907, 909 (2022).

39. Douek, *supra* note 3, at 584.

40. TIM WU, THE CURSE OF BIGNESS: ANTITRUST IN THE NEW GILDED AGE 23 (2018); Nikolas Guggenberger, *Essential Platforms*, 24 STAN. TECH. L. REV. 237, 246 (2021); Nikolas Guggenberger, *Moderating Monopolies*, 38 BERKELEY TECH. L.J. 119, 121 (2023).

41. *See, e.g.*, Nicolas P. Suzor, Sarah Myers West, Andrew Quodling & Jillian York, *What Do We Mean When We Talk About Transparency?: Toward Meaningful Transparency in Commercial Content Moderation*, 13 INT'L J. COMMC'N 1526, 1527–28 (2019); Ángel Díaz & Laura Hecht-Felella, *Double Standards in Social Media Content Moderation*, BRENNAN CTR. FOR JUST. (Aug. 4, 2021), https://www.brennancenter.org/our-work/research-reports/double-standards-social-media-content-moderation [https://perma.cc/86F4-9NGB].

A. The Rise, Fall, and Renaissance of the New Governance Paradigm

The new governance movement rose at the turn of the twenty-first century.[42] It proposed that the state, instead of enacting explicit rules, should establish "provisional and incomplete legislative frameworks" to "induce and facilitate problem solving by diffuse constituencies."[43] Its main goal was to design a process for the state, the private sector, and nonstate actors (especially civil society organizations) to share responsibility for defining and pursuing policy objectives.[44] Its toolkit centered the increased participation of nonstate actors, collaborative private-public rulemaking efforts, and the promotion of government-supported self-regulation.[45] Disclosure regimes, auditing requirements, and attention to companies' internal dynamics were meant to buttress that multistakeholder process and help monitor corporate compliance with the resulting agreements.[46]

New governance does not occupy the pages of administrative law scholarship any longer, but debates around platform regulation have imported an adaptation of this collaborative model. Scholars of content moderation took inspiration from fragments of the new governance framework to address the excessive power of corporations over online communications. Professor Kate Klonick's influential piece *The New Governors* not only was the beginning of an explosion of legal scholarship in the content moderation field but also gave a name to the paradigm that would dominate scholarly, legislative, and policy proposals in the years to come.[47] In a more recent piece, Professor Klonick makes the connection of her seminal work with the new governance movement explicit. In Professor Klonick's words: "[*The New Governors*'] title and framing draw from Professor Jody Freeman's work, among others, in the 'New Governance' movement that 'proposes a conception of governance as a set of negotiated relationships between public and private actors.'"[48] Indeed, as Professor Klonick identifies, new governance is the logic behind the proposals gaining

42. *See generally* Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267 (1998); Charles F. Sabel & William H. Simon, *Destabilization Rights: How Public Law Litigation Succeeds*, 117 HARV. L. REV. 1015 (2004); Charles F. Sabel & William H. Simon, *Epilogue: Accountability Without Sovereignty*, *in* LAW AND NEW GOVERNANCE IN THE EU AND THE US 395, 396 (Gráinne de Búrca & Joanne Scott eds., 2006) [hereinafter Sabel & Simon, *Epilogue*]; William H. Simon, *Solving Problems vs. Claiming Rights: The Pragmatist Challenge to Legal Liberalism*, 46 WM. & MARY L. REV. 127 (2004); Jody Freeman, *Collaborative Governance in the Administrative State*, 45 UCLA L. REV. 1 (1997); MARTHA MINOW, PARTNERS, NOT RIVALS: PRIVATIZATION AND THE PUBLIC GOOD (2002); Lobel, *supra* note 9.

43. Sabel & Simon, *Epilogue*, *supra* note 42, at 399.

44. Lobel, *supra* note 9, at 377.

45. Orly Lobel, *New Governance as Regulatory Governance*, *in* THE OXFORD HANDBOOK OF GOVERNANCE 65, 67 (David Levi-Four ed., 2012).

46. *Id.*

47. *See supra* note 7.

48. Klonick, *supra* note 8, at 350.

the most traction to check corporate power over the circulation of public discourse, although most of them take on the project only partially.[49]

A first partial application of the project is the creation of advisory bodies tasked with reining in corporate power on behalf of the common good. Salient examples include the Meta Oversight Board[50] and the numerous civil society advisory councils at TikTok,[51] Twitch,[52] and Spotify.[53] Their members are academics and legal experts, run nonprofits, or have careers in the public sector.[54] Their selection on the basis of credentials, prestige, and viewpoint diversity aims at ensuring that these advocates can adequately represent the interests of the public. These projects embody the kind of negotiated rulemaking process between corporate and civil society that the new governance movement promotes.

In other projects to regulate social media, nonstate actors play a less formalized but essential role.[55] Legislative proposals around the world focus on disclosure regimes, risk assessment and planning, and audited self-regulation.[56] They set up procedural requirements and bracket substantive questions to be negotiated between companies and advocates at a later stage. Most saliently, the European Union Digital Services Act strongly embraces this procedural focus.[57] Under the DSA, platforms must have unambiguous and publicly available information about their policies to moderate content, including algorithmic decision-making.[58] Companies must issue transparency reports detailing the enforcement of their policies and the actions they take regarding illegal content.[59] Very large online platforms (VLOPs)[60] must also assess "systemic risks"[61] and adopt mitigation measures, such as adapting their content moderation practices or recommendation systems.[62] VLOPs must also undergo yearly audits.[63] Less

49. *See id.*

50. *See generally* Klonick, *supra* note 11 (describing the creation and functioning of the Oversight Board).

51. Pappas, *supra* note 12.

52. TWITCH, *supra* note 12.

53. SPOTIFY, *supra* note 12.

54. *See, e.g.*, *Meet the Board*, OVERSIGHT BD., https://www.oversightboard.com/meet-the-board/ [https://perma.cc/8EKE-NNXN]; *Engaging our Advisory Councils*, TIKTOK, https://www.tiktok.com/transparency/en/advisory-councils/ [https://perma.cc/NHD9-K7BB] (listing members of the TikTok advisory councils).

55. *See* Husovec, *supra* note 17.

56. *See generally* Eric Goldman, *The Constitutionality of Mandating Editorial Transparency*, 73 HASTINGS L.J. 1203 (2022) (mapping some of the existing proposals to make content moderation more transparent).

57. Digital Services Act, *supra* note 16, art. 1(2).

58. *See id.* arts. 14, 15, 40.

59. *See id.* arts. 15–18.

60. *Id.* art. 33.

61. *Id.* art. 34.

62. *Id.* art. 35.

63. *Id.* art. 37.

developed proposals in the United States also emphasize transparency mandates and procedural requirements.[64]

As illustrated by the DSA, this paradigm does have an important role for the state.[65] Professor Freeman, a leading theorist of new governance, states that the extensive role of private actors in governance does "not imply a state which is weakened or in retreat."[66] In this account, the role of the state is to support the negotiation between companies and stakeholders.[67] The state might even participate, though only as another stakeholder.[68] Professor Freeman points out that through these negotiations, the state might even be able to extract commitments from corporations in areas that it would not be lawfully able to reach.[69]

Since Professor Klonick's initial work, scholars have developed increasingly expansive and sophisticated accounts of the internal operations of these firms and how law might support self-regulation.[70] Professor Evelyn Douek's account, for example, outlines how the state might buttress a transparent, accountable, and responsible firm.[71] Thus far, deployments of the model have been only partial, focusing primarily on self-regulation.[72] Scholars' proposals more closely resemble the new governance paradigm as they include a richer account of the role of the state—although not as robust as in the original formulations of new governance.

B. Ideological Foundations

The new governance project for content moderation offers a way out of two seemingly intractable dilemmas that have significantly paralyzed policy and legal interventions.

The first dilemma emerges from the extended belief that companies have displaced the state in exercising the public function of regulating speech.[73] Illustratively, Professors Richard Ashby Wilson and Molly Land worry that "[g]overnments are no longer the primary regulators of speech. Their

64. *See, e.g.*, Algorithmic Justice and Online Platform Transparency Act, S. 1896, 117th Cong. (2021).

65. *See* Douek, *supra* note 3, at 603.

66. Freeman, *supra* note 8, at 671.

67. *Id.*

68. *Id.*

69. *Id.*; *see also* Lee Breckenridge, *Nonprofit Environmental Organizations and the Restructuring of Institutions for Ecosystem Management*, 25 ECOLOGY L.Q. 692, 698 (1999).

70. Douek, *supra* note 3, at 603; Margot Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529, 1535 (2019); Rory Van Loo, *The Missing Regulatory State: Monitoring Business in an Age of Surveillance*, 72 VAND. L. REV. 1563, 1617–30 (2019).

71. Douek, *supra* note 3, at 603.

72. *See supra* note 12 and accompanying text.

73. Richard Ashby Wilson & Molly K. Land, *Hate Speech on Social Media: Towards a Context-Specific Content Moderation Policy*, 52 CONN. L. REV. 1029, 1032 (2021); Daphne Keller, *Who Do You Sue?: State and Platform Hybrid Power over Online Speech* 1, 2–3 (Hoover Inst., Aegis Paper Series, Paper No. 1902, 2019) ("[P]latforms can take on and replace traditional state functions, operating the modern equivalent of the public square or the post office, without assuming state responsibilities.").

regulatory capacity has been far outstripped by some of the largest companies in the world."[74] This view is supported by a discourse that refers to companies as governors, sovereigns, or regulators—all terms that evoke the figure of a state actor.[75] Indeed, use of the language of statehood to describe what social media companies do is ubiquitous.[76] Simultaneously, however, scholars also posit that the supposedly displaced state should not play the role of regulating speech directly, and constitutionally, it is significantly banned from doing so.[77]

The second dilemma results from the tension between two primary yet competing goals that advocates seek to achieve. On the one hand, advocates and critics argue that companies should have less power because the ability of a few corporate actors to control speech is at odds with values of representation and competition.[78] On the other hand, many also argue that companies should exercise their power more forcefully to control the spread of misinformation, terrorist propaganda, nonconsensual sexual imagery, and other forms of harmful speech.[79]

The result is a growing consensus around three premises that serve as the foundations of the contemporary content moderation paradigm: first, that because companies have state-like qualities, their decisions about how to regulate speech are legitimate only if they take into account the interests of the public;[80] second, that states cannot decide what those interests are because states cannot regulate lawful speech;[81] and third, that companies themselves must find mechanisms—usually taking inspiration from state institutions—to determine what values they must advance. The mechanisms most popularly proposed are that companies self-supervise and engage nonstate actors to represent the public within corporate structures.[82]

---

74. Wilson & Land, *supra* note 73, at 1032; *see also* Nadine Strossen, *United Nations Free Speech Standards as the Global Benchmark for Online Platforms' Hate Speech Policies*, 29 MICH. STATE INT'L L. REV. 307, 324–25 (2021).

75. Klonick, *supra* note 7, at 1603 ("governors"); Jonathan Peters, *The "Sovereigns of Cyberspace" and State Action: The First Amendment's Application—or Lack Thereof—to Third-Party Platforms*, 32 BERKELEY TECH. L.J. 989 (2018).

76. Josh Cowls, Phillipp Darius, Dominiquo Santistevan & Moritz Schramm, *Constitutional Metaphors: Facebook's "Supreme Court" and the Legitimation of Platform Governance*, 26 NEW MEDIA & SOC'Y 2448, 2449 (2022).

77. *See supra* note 3 and accompanying text.

78. *See, e.g.*, WU, *supra* note 40, at 133.

79. *See, e.g.*, Citron, *supra* note 37, at 69; Viktorya Vilk, Elodie Vialle & Matt Bailey, *No Excuse for Abuse: What Social Media Companies Can Do Now to Combat Online Harassment and Empower Users*, PEN AM. (Mar. 31, 2021), https://pen.org/report/no-excuse-for-abuse/ [https://perma.cc/3PBF-93K2]; CHANGE THE TERMS, RECOMMENDED POLICIES FOR TECH COMPANIES TO CURB HATEFUL ACTIVITIES ONLINE, https://www.changetheterms.org/the-terms [https://perma.cc/T53S-5PBM] (last visited Feb. 14, 2025).

80. Brenda Dvoskin, *Representation Without Elections: Civil Society Participation as a Remedy for the Democratic Deficits of Online Speech Governance*, 67 VILL. L. REV. 447, 450 (2022) (arguing that the use of the language of statehood to describe platforms changes the logics of legitimacy behind corporate behavior).

81. Douek, *supra* note 7, at 7.

82. *But see* Aviv Ovadya, *Towards Platform Democracy: Policymaking Beyond Corporate CEOs and Partisan Pressure*, BELFER CTR. (Oct. 18, 2021), https://www.belfercenter.org/publication/towards-platform-democracy-policymaking-beyond-corporate-ceos-and-

The new governance project for content moderation thus offers a way out of the two dilemmas. First, self-regulation promises to reorient the governance of speech toward public goals without state intervention. Second, self-regulation encourages responsibility and adherence to public goals without surrendering the opportunity to control the spread of harmful speech.

C. The Promise of Inclusion

At the heart of the new governance project lies a compelling promise: that the state can create a value-neutral trading zone where companies and affected constituencies engage in dialogue to solve problems collectively.[83] The process is flexible, revisable, transparent, inclusive, and constantly evolving toward "maximally informed, collaborative, and efficient solutions" that consider everyone's interests.[84] In the content moderation context, the project is appealing because it takes the accumulation of private power seriously without demanding direct state regulation over what we can say, hear, or read.

Assembling self-regulatory bodies characterized by diversity and expertise is the salient strategy for operationalizing the new governance promise of broad representation.[85] Initiatives like the Meta Oversight Board, companies' outreach to external stakeholders, and past efforts to engage advocates in speech governance aim to establish a diverse and inclusive process that steers speech governance toward the common good.[86] Accordingly, scholars' enthusiasm about the Meta Oversight Board comes from Meta's decision to "voluntarily divest itself of part of its power"[87] and bring in other interests to bear on content moderation decisions. The Board offers an opportunity not only to (admittedly, very few)[88] users to be heard but also to academics, advocates, and the general public to submit comments in every case that the board reviews.[89] It often cites the comments that it receives.[90] From the new governance perspective, the Board's positive

partisan-pressure [https://perma.cc/A7S8-HFS4] (proposing structures for more radical participation).

83. *See* Lobel, *supra* note 45, at 66 (listing the following pillars of the new governance paradigm: increased participation of nonstate actors, public/private collaboration, diversity and competition within the market, decentralization, integration of policy domains, noncoerciveness, adaptability and constant learning, and coordination).

84. Cohen, *supra* note 30, at 514 (describing and critically assessing the new governance paradigm).

85. Brenda Dvoskin, *Expertise and Participation in the Facebook Oversight Board: From Reason to Will*, TELECOMM. POL'Y, June 2023, at 1, 2.

86. On past efforts following this model, see *infra* Part III.

87. Klonick, *supra* note 11, at 2499.

88. Thomas Kadri, *Juridical Discourse for Platforms*, 136 HARV. L. REV. F. 163, 190 (2022).

89. Klonick, *supra* note 11, at 2418 (arguing that the Board "has great potential to set new precedent for user participation in private platforms' governance").

90. *See, e.g.*, *South Africa Slurs: Case Decision 2021-011-FB-UA*, OVERSIGHT BD. (Sep. 28, 2021), https://www.oversightboard.com/decision/FB-TYE2766G/ [https://perma.cc/DM4B-SRFW].

impact could be amplified with state regulation that mandates, for example, more transparency from companies and audits to ensure that Meta fulfills its responsibilities.[91]

The idea behind the diversity criterion is that if all viewpoints are represented, the outcome can be attributed to everyone.[92] The emphasis on the expertise of the actors involved is also aimed at ensuring that the outcomes of the decision-making process represent an objective stance equally beneficial to all.[93] The Board often relies on members' expertise in international human rights law, presenting its decisions as applications of this expert domain.[94] These two criteria sometimes conflict, but ultimately, they are deployed in the service of the same goal: demonstrating that self-regulation can advance the common good. Parts II and III scrutinize this promise.

II. THE EXCLUSIONARY EFFECTS OF THE NEW GOVERNANCE PROJECT FOR CONTENT MODERATION

Is it the case that companies and advocates can infuse speech governance with public goals that reflect a rational consensus among all affected stakeholders? Meta states that its stakeholder engagement efforts are guided by the values of inclusivity and expertise.[95] Its Oversight Board claims to derive its decisions from a legal framework that purports to reflect global values and consensus.[96] In the three case studies examined in Part III, the industry promised that self-regulatory bodies would work as a "public trustee,"[97] in the "public interest,"[98] and as "an agent of the public,"[99] stating that the regulation of the public sphere "belong[s] in the hands of the public."[100] During the creation of the Meta Oversight Board, scholars and activists insisted on safeguarding its independence from the company and

91. Indeed, the Meta Oversight Board has encountered difficulties to get information from Meta and monitor Meta's compliance with the Board's decisions and recommendations. *Oversight Board Demands More Transparency from Facebook*, OVERSIGHT BD. (Oct. 2021), https://www.oversightboard.com/news/215139350722703-oversight-board-demands-more-transparency-from-facebook/ [https://perma.cc/TT4X-9PDT].

92. *See* Sheila Jasanoff, *The Practices of Objectivity in Regulatory Science*, *in* SOCIAL KNOWLEDGE IN THE MAKING 307, 313 (Charles Camic, Neil Gross & Michèle Lamont eds., 2011).

93. Dvoskin, *supra* note 20, at 95.

94. *Id.*

95. *The Principles That Guide Meta's Stakeholder Engagement*, META TRANSPARENCY CTR. (Jan. 26, 2022), https://transparency.fb.com/policies/improving/principles-guide-our-stakeholder-engagement/ [https://perma.cc/FJ2B-XLAD].

96. David Kaye (Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression), *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, ¶ 41, U.N. Doc. A/HCR/38/35 (Apr. 6, 2018); Evelyn Aswad, *The Future of Freedom of Expression Online*, 17 DUKE L. & TECH. REV. 26, 65 (2018).

97. BENJAMIN, *supra* note 26, at 70.

98. FELDMAN, *supra* note 26, at 41 (National Board of Review); BLACK, *supra* note 25, at 39 (Hays Office); BENJAMIN, *supra* note 26, at 24 (NBC Advisory Council).

99. FELDMAN, *supra note* 26, at 41.

100. *Id.* at 157.

appointing a diverse body of experts—as if independence, diversity, and expertise could result in an institution capable of advancing the interests of the public.[101]

This aspiration is unattainable and counterproductive. This part makes three points. First, the public interest discourse that self-regulatory bodies produce is an outcome of the institutional environment in which they are embedded. When corporations face the threat of change in a beneficial legal environment, they aim at proving that self-regulation can work equally in the interest of all without state intervention. Implied in this analysis is the hypothesis that another institutional context would encourage different forms of self-regulation.

Second, the claim to the public interest masks the exercise of power by some groups and the exclusion of others.[102] The new governance project for content moderation amplifies the power of advocates who are critical of the industry but whose normative alignments are, relative to the position of other critics, most congruent with the pursuit of corporate profit. The argument is not that advocates' participation can never lead to good substantive outcomes. However, there is no reason to believe that the advocates who are relatively friendly to the industry have epistemic credibility in deciding public policy goals. Ultimately, efforts that emerge as a counterweight to the excessive accumulation of corporate power result in elite bodies that aim at speaking on behalf of everyone.

The exclusionary effects of the new governance model are unavoidable. Sometimes companies fail to engage with specific constituencies. Meta's exclusion of sex workers and other advocates interested in the protection of sexual expression is a clear example. However, improving the composition of advisory boards or stakeholder engagement efforts would be insufficient to overcome these deficits. The process allows only for the expression of interests that can seduce the cooperation of companies.[103] Thus, the negotiated process is inherently incompatible with its own promise of inclusion of all viewpoints.

Third, the new governance's claim to objectivity, impartiality, or a rational consensus makes contestation even more challenging and pushes dissenting voices further to the margins. Critics of the participation of nonstate actors point out that it can be ineffectual, can serve as mere cover for companies' decisions, or can offer a patina of legitimacy to business as usual.[104] In the cases examined here, the inclusion of advocates does have substantive

101. Klonick, *supra* note 11, at 2488 (surveying the concerns of those consulted during the creation of the Meta Oversight Board).

102. Fraser, *supra* note 22, at 67.

103. Cohen, *supra* note 30, at 530.

104. Lucian A. Bebchuk & Roberto Tallarita, *The Illusory Promise of Stakeholder Governance*, 106 CORNELL L. REV. 91, 94 (2020); Robyn Caplan, *Networked Governance*, 24 YALE J.L. & TECH. (SPECIAL ISSUE) 541, 541, 547 (2022); Kadri, *supra* note 88, at 189; Gilad Abri & Sebastián Guidi, *From a Network to a Dilemma: The Legitimacy of Social Media*, 26 STAN. TECH. L. REV. 92 (2023).

effects: these arrangements tend toward the depoliticization of speech governance.

The following sections take each of these points in turn.

A. Institutional Terrain

The claim that self-regulatory bodies often make of being representative of the public interest is the product of the institutional environment in which these bodies are created. Understanding this legal and institutional context and how it shapes the resulting discourse of self-regulation is essential because it enables the diagnoses of promising points of intervention. As discussed below in Part IV, using law to change this context might guide self-regulation in a more encouraging direction.

Five elements constitute the institutional environment surrounding contemporary self-regulatory efforts: the introduction of a technology that has disrupted settled communication patterns and social relations for the production of meaning; the growth of corporate power in an enabling legal environment; the rise of concerns about the harms that the new technology might create; government threats to modify the legal environment; and finally, the creation of participatory initiatives with the purpose of shielding business from aggressive state intervention. Because Meta's efforts to publicize how it self-supervises and includes advocates in online speech governance are the most developed to date, this part relies on them as rich examples. Part III will present other cases that show an analogous pattern.

The introduction of the internet disrupted settled communication patterns around broadcasting and media giants.[105] At first, it redistributed power from incumbent media giants to more decentralized speakers.[106] After a short-lived period of enthusiasm, a handful of corporate actors leveraged the new technical affordances and permissive legal environment to generate profits.[107]

As examined by Professor Julie Cohen, enabling legal rules operating in the background made social media's business model possible.[108] Social media companies designed the algorithms to organize content on their platforms thanks to the legal permission to collect data from users.[109] Indeed, these algorithms rely primarily on data that they collect about users' online behavior.[110] As Professor Cohen argues, once companies have appropriated users' data, the law protects these troves of data through contract law,

105. Benkler, *supra* note 2, at 19.

106. *See generally* ZEYNEP TUFEKCI, TWITTER AND TEARGAS: THE POWER AND FRAGILITY OF NETWORKED PROTEST (2017) (examining how individual protestors used Twitter to communicate and organize political movements).

107. *See* Benkler, *supra* note 2, at 19–20; LESSIG, *supra* note 2, at 317.

108. COHEN, *supra* note 23, at 51.

109. *Id.*

110. ARVIND NARAYANAN, KNIGHT FIRST AMEND. INST., UNDERSTANDING SOCIAL MEDIA RECOMMENDATION ALGORITHMS 1 (2023), https://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms [https://perma.cc/9S4W-C7BW].

intellectual property, and trade secrecy regimes.[111] Accordingly, companies were able to grow their business shielded from competition and regulation, concentrating power over information flows.[112] In turn, social media companies became focal points in the internet architecture where advocates and governments could target their demands about how to govern online communications.[113]

Successful advocates have developed a dual engagement with social media companies. The issue of political misinformation offers an illustration of this dual engagement, but the same case could be made about gendered harassment, terrorist propaganda, and other forms of harmful speech. On the one hand, advocates identify a problem that stems from social media's systems for distributing content. Scholars and nonprofits have observed that social media increases the reach of political misinformation.[114] On the other hand, at least some advocates have found an ally in social media firms. In the case of political misinformation and other types of lawful speech, corporate actors can offer responses that the state cannot, such as fact-checking posts, reducing posts' visibility, or simply deleting them. That is, companies can—partially—address the problem that they create by moderating the flows of information that they encourage in the first place.[115] From the advocates' perspective, this offers an opportunity to participate directly in governance.[116] Accordingly, it might even be strategic for them to focus their attention on demands that companies can address.[117] For example, advocates can partner with companies to moderate misinformation—an intervention that is more immediate and concrete than other more demanding projects such as reforming political parties or legacy media outlets in the United States.[118]

Advocates who pursue this strategy can mobilize reputational and legal risks that incentivize companies to seek out a shield against those threats, while offering them a partnership that can function as that shield. In this sense, it is essential for the continuation of these partnerships that the threats of hostile legal action do not materialize. Although at first glance it seems that the government's main role is to threaten the industry with hostile action, it is just as important that the government, in fact, is not willing or able to

111. COHEN, *supra* note 23, at 45.

112. *Id.*

113. Kadri, *supra* note 88, at 192; *see also* LESSIG, *supra* note 2 (anticipating this development).

114. *See* Anthony Nadler, Matthew Crain & Joan Donovan, *Weaponizing the Digital Influence Machine*, DATA & SOC'Y (Oct. 17, 2018), https://datasociety.net/library/weaponizing-the-digital-influence-machine/ [https://perma.cc/744P-MGQR]; Robert Faris & Joan Donovan, *The Future of Platform Power: Quarantining Misinformation*, 32 J. DEMOCRACY, no. 3, 2021, at 152.

115. COHEN, *supra* note 23, at 100.

116. *See* Caplan, *supra* note 13, at 3460–62.

117. Cohen, *supra* note 30, at 506 (arguing that the interests of advocates are constituted by the terms of the engagement process).

118. *See, e.g.*, BENKLER, FARIS & ROBERTS, *supra* note 1, at 353.

control the industry effectively. Self-regulatory efforts thrive in anticipation of unrealized state intervention.[119]

Importantly, the argument is not that self-regulation is necessarily successful in stalling government regulation. Rather, it is that managing the risk of change in a favorable legal environment is a primary factor driving how companies design self-regulatory efforts.[120] As an example, the Meta Oversight Board was created in a context of growing legislative pressure. Congressional hearings had become routine, and lawmakers were regularly introducing bills in the United States and abroad.[121] The members of the Board have themselves referenced their work as a reason to stall governmental speech regulation.[122]

Ultimately, companies include advocates to show that they can advance public goals without changes in the legal environment. As the next section examines, this incentive structure shapes which advocates have a better chance at securing a foothold in governance institutions.

B. Bad Process: The Unattainable Promise of Representing Everyone

Despite the promise of inclusive stakeholder participation, governance by consensus without exclusion is conceptually inconsistent with how representation works in the institutional context just described.[123] This section documents such exclusionary effects through three examples, from the most obvious to the most subtle: Meta's work with advocates to write, first, its nudity policies and, second, its hate speech policies and, finally, Meta's exercise of power in defining the terms of the consultations with stakeholders and the ambit of authority of its Oversight Board.

Overall, companies' desire to manage reputational and legal risks leads them to engage with advocates who share three characteristics. First, they are critical of the industry. This element is key for the engagement to be perceived as a credible counterweight to corporate power. Second, among the multiple criticisms that the industry faces, these actors tend to represent the kind of criticism that is least threatening to the pursuit of profit. Third,

119. Kadri, *supra* note 88, at 192.

120. *See id.*; Tarleton Gillespie, *Platforms Are Not Intermediaries*, 2 GEO. L. TECH. REV. 198, 201 (2018); Julie E. Cohen, *Infrastructuring the Digital Public Sphere*, 25 YALE J.L. & TECH. (SPECIAL ISSUE) 1, 28 (2023) ("[T]he dominant platforms do not want to see increasingly vociferous public critiques translated into more effective, externally imposed regulation on their infrastructuring strategies, so they prefer to speak of content removal or 'moderation' as requiring localized, post hoc adjustments to community standards and reporting policies.").

121. Douek, *supra* note 3, at 584.

122. Issie Lapowsky, *How Facebook's Oversight Board Could Rewrite the Rules of the Entire Internet*, PROTOCOL (May 6, 2020), https://web.archive.org/web/20200508013322/https://www.protocol.com/facebook-oversight-board-rules-of-the-internet [https://perma.cc/226C-TKDQ] (Chair Catalina Botero Marino stated that the "best way to . . . prevent the adoption of harmful regulations by states is for companies, in particular for the major platforms, to self-regulate.").

123. *See supra* Part II.A.

successful advocates are willing to partner with the industry because access to control points is relevant for the execution of their advocacy goals.

1. Congruent Interests: The Case of Big Tech Feminism

Tech companies hire feminists, have teams dedicated to gender and sex issues, and regularly engage feminist advocates.[124] However, not all feminists have been equally successful in getting their demands across.[125] Advocates concerned with privacy, removal of nonconsensual nudity, and combating online harassment have seen many of their proposals implemented.[126] In contrast, feminists concerned with the protection of sexual expression, speech about sex, and sex workers' participation in online platforms are a marginalized constituency.[127] These dynamics show that advocates whose projects are relatively more compatible with the industry's interests have a better chance at governing.[128]

The demand for stronger protections of sexual privacy fits better with companies' aversion to sexual expression.[129] All major social media companies now prohibit nonconsensual sexual imagery, and some have launched innovative programs to help combat this form of abuse.[130] This does not mean that advocates have not had to fight fiercely to get companies to take the nonconsensual distribution of sexual materials seriously.[131] Their important triumphs are the product of years-long campaigns. Ultimately, however, mainstream platforms can accommodate the interest in removing nonconsensual sexual images because they have an interest in removing sexual expression generally, following advertisers' preferences and perceived legal risks.[132]

---

124. DANIELLE KEATS CITRON, THE FIGHT FOR PRIVACY: PROTECTING DIGNITY, IDENTITY, AND LOVE IN THE DIGITAL AGE 174–76 (2022).

125. *See* Brenda Dvoskin, *Speaking Back to Sexual Privacy Invasions*, 99 WASH. L. REV. 59, 68–69 (2024).

126. CITRON, *supra* note 124, at 175.

127. *See, e.g.*, Kendra Albert, *Five Reflections from Four Years of FOSTA/SESTA*, 40 CARDOZO ARTS & ENT. L.J. 413, 414 (2022); Carolina Are, *The Shadowban Cycle: An Autoethnography of Pole Dancing, Nudity and Censorship on Instagram*, 22 FEMINIST MEDIA STUD. 2002, 2003 (2021).

128. *See* JANET HALLEY, PRABHA KOTISWARWAN, RACHEL REBOUCHÉ & HILA SHAMIR, GOVERNANCE FEMINISM: AN INTRODUCTION (2018) (exploring analogous dynamics in other contexts).

129. *See* JILLIAN C. YORK, SILICON VALUES: THE FUTURE OF FREE SPEECH UNDER SURVEILLANCE CAPITALISM 145 (2020).

130. *See How to Report Image Abuse on Social Media*, C.A. GOLDBERG, https://www.cagoldberglaw.com/resources/how-to-report-image-abuse-on-social-media/ [https://perma.cc/Q5B7-89G3] (reviewing the reporting mechanisms for major social media platforms, search engines, dating sites, and porn sites).

131. CITRON, *supra* note 124, at 175.

132. *See generally* YORK, *supra* note 129; *Advertiser-Friendly Content Guidelines*, YOUTUBE HELP, https://support.google.com/youtube/answer/6162278?hl=en [https://perma.cc/TM62-UA8T] (last visited Feb. 14, 2025) (outlining what kind of content is suitable for advertisers and excluding "[c]ontent that features highly sexualized content"); Albert, *supra* note 127.

In contrast, the endeavor to convince Meta to protect consensual sexual expression has so far failed.[133] Advocates have won narrow carve outs to nudity bans for content that can be clearly defined as nonsexual (for example, breastfeeding and health-related content).[134] However, even the Meta Oversight Board, which is allegedly committed to protecting speech that international law protects, has been tolerant of nudity bans.[135]

Not only do advocates have different levels of success, but Meta has also harnessed the success of privacy advocates to wield against advocates for sexual expression. These two groups are not at odds with each other. One group is concerned with the prohibition of nonconsensual distribution of sexual imagery, whereas the other group works primarily on the protection of the consensual distribution of sexual content. However, Meta justifies its broad bans on consensual sexual expression as a necessary measure to protect people's privacy.[136] Meta frames its policies as responsive to advocates' demands by invoking its work with some of them while occluding its dismissal of other groups.

Advocates' strategic behavior is not the problem. Indeed, the role of advocacy groups is to try to influence power, and the victories of privacy advocates have meaningfully improved many lives. The point is that how adjustable a normative project is to the corporate pursuit of profit is an inadequate criterion for the selection of advocates if the goal is to infuse social media governance with goals that reflect collective agreements.

2. Participatory Governance in Asymmetrical Contexts: The Case of Hate Speech Rules

Meta's process of writing its hate speech rules illustrates how companies engage with advocates in contexts of asymmetric lobbying.[137] If the objective is for advocates to overcome the marginalization of the public from speech governance, this mechanism is disappointing: it prioritizes well-organized groups and ignores other constituencies with conflicting views. This example also shows that advocates who can leverage the central control that platforms wield are more successful in cogoverning those spaces than groups that aim at contesting this order of the public sphere.

---

133. Kendra Albert, *Imagine a Community: Obscenity's History and Moderating Speech Online*, 25 YALE J.L. & TECH. (SPECIAL ISSUE) 59, 74 (2023).

134. TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA 141 (2018).

135. *Breast Cancer Symptoms and Nudity: Case Decision 2020-004-IG-UA*, OVERSIGHT BD. (Jan. 28, 2021), https://www.oversightboard.com/decision/IG-7THR3SI1/ [https://perma.cc/K4CN-APPB].

136. *Adult Nudity and Sexual Activity*, META, https://transparency.fb.com/policies/community-standards/adult-nudity-sexual-activity/ [https://perma.cc/JMR8-L7WK] ("[W]e default to removing sexual imagery to prevent the sharing of non-consensual or underage content.").

137. Dvoskin, *supra* note 80, at 478.

Racial justice and civil rights organizations have strongly campaigned for more restrictions on racist and discriminatory speech.[138] Companies have received little pushback to protect these forms of harmful but lawful expressions.[139] This asymmetric lobbying could be attributed to a change in cultural norms.[140] On a closer look, advocates for stricter regulation of hate speech have been more effective in articulating their normative preferences and organizing their lobbying efforts than their counterparts.[141] Advocates for robust speech protections have been paralyzed by their dual commitment to free speech and market freedom, being unable to formulate clear demands to corporate actors.[142] In contrast, advocates who favor both state and nonstate restrictions on harmful speech have been able to seize private control over online communications.[143] This asymmetry has contributed to the evolution of content moderation rules toward increasingly expanded definitions of the content that platforms disallow.[144]

Comparing the reaction of Jonathan Greenblatt, Chief Executive Officer of the Anti-Defamation League, to YouTube's decision to ban Holocaust denial with the reaction of Jennifer Granick, a member of the American Civil Liberties Union (ACLU), illustrates this asymmetry. Although the former advised Meta to follow YouTube's example and adopt the same ban, the latter cautioned that "YouTube will make mistakes and over-censor," but "as a private company is well within its rights."[145] This example captures the core of the asymmetry: although Greenblatt has a specific rule that he would urge all services to adopt, the ACLU does not urge the adoption of any specific balance of values.[146]

Advocates for restrictions on lawful but harmful speech share the three characteristics described above. First, they are strong critics of content moderation companies. For example, the Center for American Progress, Color Of Change, The Free Press, MediaJustice, Muslim Advocates, the National Hispanic Media Coalition, the Southern Poverty Law Center, and the Lawyers' Committee for Civil Rights Under Law have formed a coalition

138. *See, e.g.*, *More Than 500 Companies Join Facebook Ad Boycott "Stop Hate for Profit,"* BUS. & HUM. RTS. CTR. (July 1, 2020), https://www.business-humanrights.org/en/latest-news/more-than-500-companies-join-facebook-ad-boycott-stop-hate-for-profit/ [https://perma.cc/G5PU-QK7K].

139. Feldman, *supra* note 3, at 101.

140. *See, e.g.*, Evelyn Douek, *Governing Online Speech: From "Posts-as-Trumps" to Proportionality and Probability*, 121 COLUM. L. REV. 759, 778 (2021) (describing a "cultural reevaluation" of how social media companies should moderate speech).

141. Dvoskin, *supra* note 80, at 479.

142. *See* Hannah Bloch-Wehba, *The Rise, Fall, and Rise of Cyber Civil Libertarianism*, *in* FEMINIST CYBERLAW 153, 154 (Mega Leta Jones & Amanda Levendowski, eds., 2024).

143. Dvoskin, *supra* note 80, at 478.

144. *Id.* at 463; Bloch-Wehba, *supra* note 142, at 159.

145. Paresh Dave, *YouTube Reversal Bans Holocaust Hoaxers, Stops Pay for Borderline Creators*, REUTERS (June 5, 2019, 5:34 PM), https://www.reuters.com/article/us-alphabet-youtube-hatespeech/youtube-reversal-bans-holocaust-hoaxers-stops-pay-for-borderline-creators-idUSKCN1T623X [https://perma.cc/4EPD-VDYX].

146. Dvoskin, *supra* note 80, at 478.

called Change the Terms.[147] The coalition's mission is to denounce large social media platforms' hosting of various forms of discriminatory content.[148] Besides being vocal critics, they have been able to generate reputational and legal risks for the companies, especially Meta, having mobilized advertisers to boycott the company.[149] Between 2018 and 2020, Meta conducted a civil rights audit at the behest of these organizations.[150] The final report publicly showed Meta's failures to combat discrimination.[151] The concerns raised by these organizations have been frequently echoed by members of Congress.[152]

Second, the claims made by these advocates pose the optimal level of risk for the company. These advocates have created a risk that Meta is interested in managing, but they express concerns that Meta has been able to—at least partially—address. Their demands have centered primarily on adjusting the rules that determine what content is or is not allowed on the platform. They do not impact how the company collects and processes data from and about users to distribute speech and advertising. Thus, it has been possible for the company to satisfy these demands by adjusting rules that might have a large impact on the digital speech environment but only a marginal impact on its business model.

Finally, access to online speech governance is a valuable asset that companies can offer to these groups. The announcement that launched Change the Terms' proposal for content moderation rules stated that "[b]ecause internet tools are largely owned and managed by the private sector and not government, *these corporations must be part of the solution* to address the promulgation of hateful activities online."[153] In fact, private governance provides a unique opportunity because the companies can ban constitutionally protected speech.[154] In other words, the advocates with the most influence are those whose projects can benefit from the control that gatekeepers wield. By contrast, the project of civil-libertarian advocates has a more ambivalent relation to corporate gatekeeping power and therefore

147. *About*, CHANGE THE TERMS, https://www.changetheterms.org/about/ [https://perma.cc/4UBC-GABA].

148. *Id.*

149. *See, e.g.*, *WFA and Platforms Make Major Progress to Address Harmful Content*, WORLD FED'N OF ADVERTISERS (Sep. 23, 2020), https://web.archive.org/web/20200925011257/https:/wfanet.org/knowledge/item/2020/09/23/WFA-and-platforms-make-major-progress-to-address-harmful-content [https://perma.cc/P5QN-J8YG]; Susan Wojcicki, *Letter from Susan: Our 2021 Priorities*, YOUTUBE OFF. BLOG (Jan. 26, 2021), https://blog.youtube/inside-youtube/letter-from-susan-our-2021-priorities/ [https://perma.cc/27US-6KZQ] ("Our policies are designed to protect our YouTube community against abuse and bad actors and also to make sure we are able to keep advertisers coming back to YouTube . . . .").

150. LAURA W. MURPHY, FACEBOOK'S CIVIL RIGHTS AUDIT—FINAL REPORT 1, 3 (2020), https://about.fb.com/wp-content/uploads/2020/07/Civil-Rights-Audit-Final-Report.pdf [https://perma.cc/Y5ZS-ND2K].

151. *Id.*

152. *Id.*

153. Center for American Progress et al., *supra* note 79 (emphasis added).

154. Dvoskin, *supra* note 80, at 478.

offers companies no unambiguous path toward communicating to the public that they are advancing the common good.

Again, the point is not that one side of the debate within civil society is correct and the other is not. The Anti-Defamation League and the ACLU both pursue a vision of speech regulation that they believe in earnestly, and both hold a legitimate view of how the public sphere in a democratic society could function. The point is that on deeply contested questions of speech governance, the model built on corporate engagement with advocates suffers a structural bias toward the side that most aligns with corporate profits, irrespective of its theoretical, democratic, or moral strength.

3. Defining the Terms of Engagement

Engaging advocates—even with the purpose of reallocating corporate power to inclusive, diverse, and balanced multistakeholder fora—is a practice of power with distributional effects.[155] For the process to work, advocates must espouse interests likely to win the cooperation of companies. The governance structure requires that those who participate frame their projects in ways that fit the terms of the engagement, which are protective of the existing order of the public sphere.[156]

When Meta reached out to dozens of academics, civil society groups, and international organizations to discuss whether images of Zwarte Piet—the Dutch character often portrayed in blackface—should be allowed on the platform, the terms of the engagement had already selected for what interests would be represented at (and excluded from) the table.[157] No matter how inclusive these engagement processes are, they allow expression of only a predefined set of interests (in favor or against a content moderation rule). Both results are acceptable from the company's perspective. From advocates' perspectives, some might be interested in the discussion, others might have preferred to focus on procedural aspects, and others might have wanted Meta to change the design of its interface, how users can moderate groups, or any other of a number of options. Even more, some advocates who work against racist speech might have reallocated some of their resources to think about how Meta should deal with this particular problem given the opportunity to bring about change.[158] In all these different ways, these processes select for and reconfigure advocates' interests.

The selection of the members of the Meta Oversight Board also exemplifies how the terms of the engagement produces exclusionary effects even in initiatives prioritizing diversity. When the Board was created, Meta

155. Cohen, *supra* note 30, at 530.

156. *Id.* at 539–40 (describing the efforts to empower people to negotiate as a practice of power that itself constitutes participants' interests).

157. *Facebook Bans Zwarte Piet Blackface Images and Racist Depictions of Jews*, EURONEWS (Aug. 12, 2020), https://www.euronews.com/2020/08/12/facebook-bans-zwarte-piet-blackface-images-and-racist-depictions-of-jews [https://perma.cc/7WPA-3F7T].

158. A fundamental aspect not discussed here is the influence that funders and the way they measure impact have on advocates' agendas and their engagement with corporate processes.

chose members that simultaneously represent diverse constituencies and bring expert knowledge. According to Professor Noah Feldman's original proposal, the Board's members should "enjoy national or international public reputations within their professions. They could be of various ages and career stages, but they should bring prestige to the Facebook Supreme Court at least as much as they derive prestige from it."[159] Meta has been highly successful in meeting this standard. Members include former judges, former United Nations and Inter-American Commission on Human Rights special rapporteurs, civil society leaders, and university professors.[160] Diversity has also been key to choosing the members of the Board. Even though recurring to expertise has been the company's dominant strategy in fostering the authority of the Board, maintaining a balanced composition has also been a priority.[161]

Taken together, the prestige and diversity of the Board's members are important to building a credible claim that they can act as counterweight to Meta's power. Furthermore, many of them had voiced important critiques of content moderation before their appointment. For example, Board members, Professors Evelyn Aswad and Nicolas Suzor, had been vocal critics of Meta's content moderation policies before being appointed.[162] Indeed, at the time of their appointment, former special rapporteur for freedom of expression Professor David Kaye said of Professor Suzor that he "wrote a key book on social media with a title ('Lawless') that *makes you wonder how he got through the selection gauntlet*."[163] Further, the members see themselves as strong counterweights to the company's power.[164]

However, when the kind of concerns that the members bring to the table are examined, it is easy to understand why they were selected. They focus on concerns that the industry can easily absorb without putting its business interests on the line. In particular, Professors Aswad and Suzor have been strong critics of social media companies for not aligning their content moderation rules with international human rights law.[165] As discussed in more detail below, making recommendations on how to align content moderation with human rights law has been the project of the Oversight Board. Importantly, the company might be relatively open to adopting

159. Feldman, *supra* note 3, at 109 (emphasis omitted).

160. *See* OVERSIGHT BOARD, *supra* note 54.

161. Klonick, *supra* note 11, at 2457.

162. *See* Evelyn Aswad, *supra* note 96, at 28. *See generally* NICOLAS P. SUZOR, LAWLESS: THE SECRET RULES THAT GOVERN OUR DIGITAL LIVES (2019).

163. David Kaye, *The Republic of Facebook*, JUST SECURITY (May 6, 2020) (emphasis added), https://www.justsecurity.org/70035/the-republic-of-facebook/ [https://perma.cc/5Z7J-3PYM].

164. Elizabeth Culliford, *Facebook Names First Members of Oversight Board That Can Overrule Zuckerberg*, REUTERS (May 6, 2020, 6:58 PM), https://www.reuters.com/article/us-facebook-oversight/facebook-names-first-members-of-oversight-board-that-can-overrule-zuckerberg-idUSKBN22I2LQ [https://perma.cc/2EAL-Q3CE] (Professor Suzor stated "[w]e're not working for Facebook, we're trying to pressure Facebook to improve its policies and its processes to better respect human rights. That's the job.").

165. Aswad, *supra* note 96, at 35; SUZOR, *supra* note 162.

slightly amended content moderation policies—especially if doing so diffuses public concerns about content moderation.

The internal logic of the Board selects for concerns that can accommodate existing corporate structures. Even if more radical critics of Meta were invited to join the Board, to participate in the endeavor and secure Meta's cooperation, they would need to adapt their interests to the scope of possible agreements. The Board's ambit of authority is narrow.[166] It can make recommendations to Meta and has used that feature creatively to broaden its impact, for example, by asking Meta to assess the enforcement of its content moderation rules in the context of the Israeli-Palestinian conflict.[167] Yet, it has no power to oversee the company's privacy settings, recommendation systems, interface design, ad dashboard, or any other aspect germane to the company's advertising business.[168]

In short, multistakeholder fora or arrangements like the Meta Oversight Board prioritize viewpoints that do not threaten—and that usually benefit from—the market structure or legal conditions supporting companies' dominant positions. It is impossible to predict whether the work of the Board and analogous bodies will result in changes that gain the support of large parts of the public or whether it will defuse critical voices while producing insignificant change.[169] Nonetheless, this governance model empowers a specific set of interests, no matter how diverse the group might be.

C. Bad Outcomes: A Counterproductive Claim to the Public Interest

When self-regulatory bodies make decisions, they claim to represent the public interest or an objective stance. This type of discourse occludes their function as a vehicle for the exercise of power by some. Instead of acknowledging the need to create institutional opportunities for excluded groups to contest the decisions of governing elites, this claim affirms that self-regulation produces outcomes in the interest of all. This section explores three related consequences: the masking of the power of elites over public discourse, the hampering of productive disagreement, and the subversion of advocates' role.

1. Masking Power

An examination of the doctrinal work of the Meta Oversight Board shows that the Board's performativity as an objective decision-maker hides its

166. The Board can only decide whether to take down or keep up specific posts and make policy recommendations that are not binding for Meta. META, OVERSIGHT BOARD CHARTER, art. 2 (2019), https://about.fb.com/wp-content/uploads/2019/09/oversight_board_charter.pdf [https://perma.cc/3ELG-DU7H].

167. *Shared Al Jazeera Post: Case Decision 2021-009-FB-UA*, OVERSIGHT BD. (Sep. 14, 2021), https://www.oversightboard.com/decision/FB-P93JPX02/ [https://perma.cc/9X9S-53F3].

168. Douek, *supra* note 29, at 387.

169. *See generally* Cohen, *supra* note 30.

policy decisions. This should be no surprise to legal scholars since courts do analogous work.[170]

Professor Sheila Jasanoff explains that experts attain authority through "a highly sought-after and hard-won epistemic achievement, namely objectivity."[171] In the case of self-regulatory bodies, an objective account would be one that provides a rational view of the public interest and represents a standpoint equally in the interest of all.[172] As described in Part II.B, diversity is often a tool used to convey that bodies represent what Professor Jasanoff calls a "view from everywhere."[173] Following the Habermasian ideal, if all viewpoints are represented, the decisions might be attributed to everyone.[174] The previous section argued that self-regulatory bodies' claim of offering a view from everywhere is an illusion. Another common mechanism that experts rely on is what Professor Jasanoff, borrowing from Professor Thomas Nagel, describes as a "view from nowhere."[175] A view from nowhere conveys to the public that decision-makers act based not on their subjective viewpoint or capricious motives but on some exogenous principles or existing agreements that ground and constrain their actions.[176]

To cultivate the credibility of its view from nowhere, the Meta Oversight Board attempts to act as a decision-maker that implements exogenous principles with global legitimacy and consent—namely international human rights law.[177] However, significant areas of IHRL are vague and even contradictory.[178] In addition, the Board has done a great deal of interpretative work to introduce indeterminacy in IHRL.[179] This work allows the Board to use IHRL language to conceal its policy preferences as objective applications of exogenous principles.[180]

Proving that IHRL does not actually produce objective outcomes, the Board has allowed Meta to adopt rules that, in its own view, directly

170. *See generally* Oliver Holmes, *The Path of the Law*, 10 HARV. L. REV. 457 (1897); John Dewey, *Logical Method and Law*, 10 CORNELL L.Q. 17 (1924) (showing that legal reasoning conceals policy decisions).

171. Jasanoff, *supra* note 92, at 308.

172. *Id.*

173. *Id.* at 314.

174. *See* JÜRGEN HABERMAS, MORAL CONSCIOUSNESS AND COMMUNICATIVE ACTION 43 (Christian Lenhardt & Shierry Weber Nicholsen trans., 1990) (1983).

175. Jasanoff, *supra* note 92, at 309. *See generally* THOMAS NAGEL, THE VIEW FROM NOWHERE (1986).

176. Jasanoff, *supra* note 92, at 309; *see also* Klonick, *supra* note 11, at 2418 (describing the creation of the Meta Oversight Board and highlighting that "[a] critical piece of establishing the independence of an adjudicatory body is ensuring that members are without, or are willing to set aside, personal bias on matters before them").

177. Dvoskin, *supra* note 20, at 112.

178. Evelyn Douek, *The Limits of International Law in Content Moderation*, 6 U.C. IRVINE J. INT'L, TRANSNAT'L & COMPAR. L. 37, 53, 56 (2021).

179. Dvoskin, *supra* note 20, at 109.

180. *Id.* at 108.

contradict what IHRL prescribes.[181] Even in these cases, the Board has justified these rules as an appropriate implementation of human rights law. A decision reviewing Meta's prohibition on content depicting blackface is instructive.[182] On this occasion, the Board considered that "international human rights law would not allow a state to impose a general prohibition on blackface through criminal or civil sanctions" but simultaneously concluded that "[Meta] followed international guidance and met its human rights responsibilities in this case."[183] The Board had to explain why a state-issued ban on blackface would violate international law but Meta was nonetheless meeting its human rights responsibilities when issuing the exact same ban.[184]

The Board gave four reasons to justify Meta's ban as a correct application of IHRL: namely, that United Nations experts and other authorities have determined that blackface creates objective harm,[185] that Meta has a responsibility to promote equality,[186] that the intent of the speaker on social media is hard to evaluate,[187] and that the accumulation of online hate speech can create a discriminatory and degrading environment.[188] Regarding the first rationale, the harm that blackface creates is independent of which authority regulates it. On the second rationale, it is clear that the state's duty to promote equality is stronger than Meta's responsibility. The two last rationales point to differences between online and offline speech, not to differences between states and companies.

The Board relied heavily on the specificities of online speech that would, in its view, make blackface content more harmful than in the offline context. Even if this is the case, it is incoherent to argue that online speech merits new normative answers (hinting that IHRL stands at an inflection point) and assert that states would breach international law if they banned images depicting people in blackface (indicating that IHRL prescribes stable rules in this area).[189]

The Board could have argued that in the new online context, IHRL is insufficient to answer all questions and normative decisions that need to be made. However, this option would have been too costly. It would open the door to states' departing from IHRL and regulating, in the online context, speech that IHRL traditionally protects. This option likely conflicts with the views of at least some of the Board's members.[190] More importantly, the

181. *Depiction of Zwarte Piet: Case Decision 2021-002-FB-UA*, OVERSIGHT BD. (Apr. 13, 2021), https://www.oversightboard.com/decision/FB-S6NRTDAJ/ [https://perma.cc/EF4R-S7MM]; *South Africa Slurs*, *supra* note 90.

182. Dvoskin, *supra* note 20, at 119 (analyzing this decision).

183. *Depiction of Zwarte Piet*, *supra* note 181.

184. *Id.*

185. *See id.* at 16.

186. *Id.*

187. *See id.*

188. *Id.*

189. *See* Dvoskin, *supra* note 20, at 121.

190. Catalina Botero Marino (Special Rapporteur on Freedom of Expression, Inter-American Commission of Human Rights [Inter-Am. Comm'n H.R.]), *Annual Report of*

Board would lose the highly sought after legitimizing force of its practice of appealing to IHRL.

The decision illustrates that even the most sophisticated efforts to implement global agreements do not advance a rational consensus among all. Even within the possible outcomes that self-regulatory bodies allow, their decisions do not represent collective agreements and have distributive effects. The evaluation of this outcome depends on two factors. First, one might be more or less supportive of it depending on whether one agrees or not with the views of those who prevail. Second, this outcome might be the best alternative in the absence of institutional options to create governance structures for more effective sharing of power. Part IV resumes this discussion.

2. Making Contestation Harder and Renouncing Politics

As examined above, the members of self-regulatory initiatives portray their work as an expression of the public interest with democratic buy-in. This discourse makes contestation harder. Because decisions are framed as objective instead of political preferences, outsiders' challenges become more burdensome. Relatedly, these initiatives pursue a vision of advocacy where advocates seek out a rational consensus with the industry and among themselves instead of engaging in political struggle. Both elements hinder the contestation of the hegemonic order of the public sphere.

The Meta Oversight Board once again provides a good example. The Board argues that the United Nations Guiding Principles on Business and Human Rights (UNGPs) set the expectation that the Board will apply IHRL. That is, the Board argues not only that its decisions are an objective implementation of IHRL but that the application of the IHRL framework itself stems from an exogenous instrument.

However, the UNGPs do not suggest that companies should use IHRL as their own default rules to moderate content in any obvious way.[191] The UNGPs set the expectation that business enterprises will respect human rights recognized by international law regardless of the domestic law in the jurisdictions where they operate.[192] Originally, the principles were a response to companies' failure to respect human rights under the excuse that local law did not compel them to do so.[193] However, it is unclear which human right the Meta Oversight Board should safeguard. At first sight, the most relevant human right in this sector is the right to freedom of expression.

---

the Special Rapporteur on Freedom of Expression, 493, OEA/Ser.L/V/II.149 Doc. 50 (Dec. 31, 2013).

191. Dvoskin, *supra* note 20, at 104.

192. U.N. Special Representative of the Secretary-General, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework*, 13, U.N. Doc. A/HRC/17/31 (Mar. 21, 2011).

193. *See* John Gerard Ruggie, *Incorporating Human Rights: Lessons Learned, and Next Steps*, *in* BUSINESS AND HUMAN RIGHTS: FROM PRINCIPLES TO PRACTICE 64, 65 (Dorothée Baumann-Pauly & Justine Nolan eds., 2016).

Accordingly, scholars argue that Meta should use IHRL to inform its own editorial policies to meet its human rights responsibilities.[194] However, IHRL provides strong protections for editorial freedom and does not recognize a right to speak on social media platforms without being subjected to the editorial discretion of the medium owner. Then, it is unclear why Meta would need to make editorial decisions following international human rights norms.[195]

The Board's assertion that it must apply IHRL positions it as making "correct" judgments instead of contestable decisions. Importantly, the Board has been more protective of dissent than Meta had been in the past, and the language of IHRL has been employed to support this normative orientation.[196] However, this rhetorical device makes it more burdensome for users appealing decisions and submitting comments to propose rules that diverge from IHRL. Similarly, for other companies making different decisions, it is costlier to justify a different set of rules, as the Board's recommendations appear to be mandated by some higher authority. Indeed, the Board has shown interest in expanding its authority over other companies, thereby promoting more concentration in decision-making power.[197]

3. Subverting Advocates' Role

The Board's claim of objectivity also prevents it from living up to its full potential. Its work to frame its decisions in IHRL language hampers the Board in fulfilling its goals of providing public justifications of how it balances different interests, proposing normative projects, engaging with other proposals, and subjecting its own decisions to disagreement.[198]

The decision in the case regarding the suspension of then-former President Donald J. Trump's account shows how the discourse of objectivity hinders its potential.[199] When the Board made a decision about whether Meta had been right or wrong in suspending Trump from using its platforms, Professor Douek criticized the Board for "resolv[ing] essentially nothing."[200] The Board upheld Meta's initial action but asked Meta to write rules to govern when users should be excluded from the platform, which in turn the Board would review in the future. Professor Douek fairly considered the decision to be "too modest," echoing other critics.[201] However, this critique assumes

194. *See* Aswad, *supra* note 96, at 64–70; Strossen, *supra* note 74, at 314–17.

195. *See* EMILY B. LAIDLAW, REGULATING SPEECH IN CYBERSPACE: GATEKEEPERS, HUMAN RIGHTS AND CORPORATE RESPONSIBILITY 91–93 (2015).

196. Strossen, *supra* note 74, at 331.

197. *See* Kadri, *supra* note 88, at 164.

198. Feldman, *supra* note 3.

199. *See Former President Trump's Suspension: Case Decision 2021-001-FB-FBR*, OVERSIGHT BD. (May 5, 2021), https://www.oversightboard.com/decision/FB-691QAMHJ/ [https://perma.cc/JXD5-SWBM].

200. Evelyn Douek, *Somebody Has to Do It*, ATLANTIC (May 6, 2021), https://www.theatlantic.com/ideas/archive/2021/05/facebook-oversight-board-trump-democracy/618825/ [https://perma.cc/S5HM-WNA3]; *see also* Douek, *supra* note 29, at 405–06.

201. Douek, *Somebody Has to Do It*, *supra* note 200; Mark MacCarthy, *The Facebook Oversight Board's Failed Decision Distracts from Lasting Social Media Regulation*,

that the Board could have readily done otherwise. This "modesty" in difficult cases is by design. As Professor Douek highlights, the Board "is *not* actually incentivized to weigh in on intractable normative and moral debates, despite this being its ostensible purpose."[202] The Board has done a great deal of work to portray itself as an objective body drawing on its legal expertise as opposed to its political commitments. In cases with less visibility, the open texture of IHRL is sufficient for the Board to conceal its power. In cases with high visibility, however, deferring the difficult questions is a reasonable strategy to maintain its authority. The Board is invested in its own success, which depends on it being perceived as an objective body capable of making nonpolitical, authoritative decisions. Lacking actual democratic representativeness, the need to maintain a court-like appearance with neither the legal authority of courts nor the historical bases of their legitimacy renders this hybrid structure weakest precisely when the political stakes of its decisions are most salient.

If the goal is to create boards that express their political commitments forcefully and participate in the public debate over how to govern speech, the legal framework must create a different institutional terrain and incentive structure.

III. BEFORE SOCIAL MEDIA

The main claim of this Article is analytical: specifically, that the collaborative process that the new governance model envisions is a practice of power that benefits the viewpoints most adaptive to the hegemonic order of the public sphere. Looking at contemporary and historical—admittedly, imperfect—deployments of this model shows that this concern is not merely hypothetical. The attempts to construct the public interest through the collaboration between companies and advocates have not worked. Looking at the interests that have historically been excluded further underscores the practical importance of this conceptual concern. In a stratified society, marginalized groups tend to suffer more severe exclusions from the governance elite.[203]

This part examines three historical case studies of advocates partnering with the industry to govern speech on behalf of the public interest under institutional conditions analogous to the ones surrounding social media today. Indeed, even though the power of social media companies is sometimes framed as unprecedented, the film and radio industries went through surprisingly similar patterns almost a century ago. The Edison Trust

BROOKINGS INST. (May 11, 2021), https://www.brookings.edu/blog/techtank/2021/05/11/the-facebook-oversight-boards-failed-decision-distracts-from-lasting-social-media-regulation/ [https://perma.cc/RR3Q-ENXD]; *Facebook Oversight Board Decision to Ban Trump Doesn't Solve Disastrous Handling of Hateful and Violent Speech*, AMNESTY INT'L (May 5, 2021), https://www.amnesty.org/en/latest/news/2021/05/facebook-oversight-board-decision-to-ban-trump-doesnt-solve-disastrous-handling-of-hateful-and-violent-speech-2/ [https://perma.cc/R4SZ-8VYD].

202. Douek, *supra* note 29, at 433 (emphasis added).

203. Fraser, *supra* 22, at 67.

(the first association of film producers), the Hollywood film studios, and the radio National Broadcasting Company accumulated vast power over films and radio.[204] These industries were at the time the most popular channels for entertainment and cultural discourse.[205] Lawmakers and advocacy groups raised two types of concerns about the content of films and radio programming.[206] The dominant position of a few actors in controlling public discourse was a source of distress.[207] In addition, advocates, religious groups, and government actors demanded that the industries use their power to curb the spread of expression considered harmful, offensive, dangerous, or immoral.[208] In response, the three industries designed advisory boards to govern the production and distribution of films and radio programming: the National Board of Review (NBR), the Hays Office (the "Office"), and the NBC Advisory Council (the "Council").[209]

These three institutions had the explicit goal of persuading lawmakers that state intervention was unnecessary because they would govern speech on behalf of the public interest.[210] Despite this promise, these governance arrangements empowered only those actors whose interests were congruent with the corporate pursuit of profit and entrenched corporate power over the production of speech. The case studies also reveal that legal reforms that were available at the time had the potential for advancing a plural, egalitarian, and contested public sphere more forcefully than the boards that were set up to represent the interests of the public.

Before proceeding, a note on method is in order. These cases are relevant because of the institutional similarities with the social media governance context. This does not mean, however, that the cases are in any way identical. The differences are abundant. Far from the complex managerial structures developed by tech giants or the elaborated Digital Services Act, the Edison Trust, Hollywood, and NBC developed more rudimentary institutions. These stories are about purely self-regulatory bodies, without state support. The next part explains why more sophisticated versions and different forms of state support for these initiatives do not challenge the institutional limitations that the case studies reveal.

In addition, the new governance school of thought emerged after the deregulatory neoliberal turn. However, this regulatory ideology has historically marked the United States' style of engagement with the law.[211] An embryonic version of the new governance school's ideological premises drove the development of the models examined here.

---

204. *See* STARR, *supra* note 25, at 307; BLACK, *supra* note 25; STERLING & KITTROSS, *supra* note 25.

205. *See* STARR, *supra* note 25, at 307; BLACK, *supra* note 25; STERLING & KITTROSS, *supra* note 25.

206. *See infra* Parts A.1, B.1, C.1.

207. *See infra* Parts A.1, B.1, C.1.

208. *See infra* Parts A.1, B.1, C.1.

209. FELDMAN, *supra* note 26; RANDALL, *supra* note 26; BENJAMIN, *supra* note 26.

210. *See, e.g.*, BENJAMIN, *supra* note 26, at 12.

211. *See* Jodi L. Short, *The Paranoid Style in Regulatory Reform*, 63 HASTINGS L.J. 633, 635–38 (2012).

The legal environment was also widely different, especially with respect to the First Amendment and the structure of the administrative state. Functionally, however, the legal environment was analogous—even if almost exactly opposite in doctrinal terms. A 1915 Supreme Court ruling shows how contradictory First Amendment doctrine was vis-à-vis its contemporary counterpart at the time.[212] In *Mutual Films Corp. v. Industrial Commission of Ohio*,[213] the Court decided that films were not covered by the First Amendment because they were entertainment, not "the press."[214] However, what matters to our understanding of how law shaped the interests of the industry is not the specifics of First Amendment doctrine but the fact that the industry was concerned about the loss of a generally beneficial legislative environment. At the time, firms were concerned about threats of federal censorship laws. Proposals to legislate privacy and personal data, regulate advertising practices, enforce antitrust action, create interoperability mandates, or pass Section 230[215] reform play analogous roles today.[216]

The administrative state has been transformed since then. The period examined here corresponds to the period before or during the incipient creation of independent agencies. Even in the case of radio, there was initially no intervention in the market structure, and as Professor Paul Starr says, radio would have "no new deal."[217] However, despite the massive size of the current administrative state, the lack of capabilities to meaningfully oversee the information economy makes it more similar to the pre–New Deal administrative state than it might appear at first sight.[218] The Edison Trust, Hollywood, and NBC evolved in an environment of constant threat of legislation that would increase the costs of conducting business, while at the same time, the state had no meaningful capabilities to regulate or monitor the industry. In that environment, advocates were tasked with organizing the production and distribution of speech on behalf of the public interest. The next sections examine the three case studies in chronological order.

A. The Case of the Edison Trust and the National Board of Review

The National Board of Review was the first attempt in the film industry to self-regulate the content of movies. The main goal driving its creation was to stop federal censorship bills, circumvent state censorship boards, and prevent other forms of hostile state action such as the closure of film theaters. In 1912, the NBR reviewed ninety-eight of every 100 movies exhibited in

212. *See* Mut. Films Corp. v. Indus. Comm'n of Ohio, 236 U.S. 230, 234 (1915).
213. 236 U.S. 230 (1915).
214. *Id.*
215. Communications Decency Act of 1996, Pub. L. No. 104-104, § 230, 110 Stat 56, 137–39 (1996).
216. *See infra* Part IV.C.
217. STARR, *supra* note 25, at 361.
218. *See* COHEN, *supra* note 23, at 170 (arguing that the administrative state, designed for the industrial economy, is ill-equipped to govern information capitalism).

the country.[219] The result was more sanitized films that abandoned the political discussion of the early days of film production, although the censorship bills that the NBR was set up to prevent would have likely been even more censorial. The NBR lost control over films when the trust of film studios lost control over its pool of patents. Then, independent producers reentered the movie market. For a short-lived period, diversity and political issues returned to the screen. The case study illustrates the incentives of the industry to partner with advocates, the kinds of advocates who gained access to governance, and the sanitizing effects of this alliance. It also shows that background legal rules—in this case, patent law—are an essential component of the speech environment and a promising site for the state to intervene. The first subsection describes the events leading to the creation of the NBR. The second subsection discusses its effects.

1. The Institutional Terrain Producing the National Board of Review

The period between 1903 and 1908 was a time of great expansion for the film industry.[220] Especially among the working class, motion pictures were the most popular form of entertainment, and they became a central medium for political critique.[221] Early movies criticized Victorian moral standards and addressed labor issues, the struggles of immigrants, and generally the lives of poor people in cities.[222] By 1907, civil society and religious groups were concerned about the content of movies.[223] They believed that the new technology had an especially powerful influence on young minds.[224] An editorial from 1907 claimed that nickelodeons were "schools of crime where murders, robberies, and holdups are illustrated."[225] It concluded that "[t]he proper thing for city authorities to do is to suppress them at once."[226]

The dissatisfaction of advocacy and Protestant groups became influential only when three conditions were met: the state was threatening the industry with hostile measures, advocates had opted for the strategy of partnering with the industry to govern films, and the Edison Trust was producing most of the

219. FELDMAN, *supra* note 26, at 8.

220. *See* LEWIS JACOBS, THE RISE OF THE AMERICAN FILM 52 (1939); STARR, *supra* note 25, at 305.

221. BLACK, *supra* note 25, at 7.

222. *See id.*

223. *Id.* at 8.

224. *See* STARR, *supra* note 25, at 305; *see also* Kathleen McCarthy, *Nickel Vice and Virtue: Movie Censorship in Chicago, 1907–1915*, 5 J. POPULAR FILM 37 (1976); BLACK, *supra* note 25, at 9, 50; LARY MAY, SCREENING OUT THE PAST 40 (1983) (A professor at the University of Kansas stated: "Pictures are more degrading than the dime novel because they represent real flesh and blood characters and import moral lessons directly through the senses.").

225. FELDMAN, *supra* note 26, at 3.

226. *Id.*; *see also* BLACK, *supra* note 25, at 9–10 (quoting a YMCA official stating that "[u]nless the law steps in and does for moving-picture shows what it has done for meat inspection and pure food, the cinematograph will continue to inject into our social order an element of degrading people").

domestic films. These conditions of threatened state action, industry concentration, and company-advocate partnerships also characterize the context of social media self-regulatory efforts.

Take first the threats coming from the state. In 1908, New York City Mayor George B. McClellan Jr., echoing the concerns of advocates, revoked every theater's license.[227] The theaters were required to apply for new licenses under new rules that did not allow them to open on Sundays and that prohibited them from showing movies that tended to "degrade or to injure the morals of the community."[228] Although the theater owners managed to secure a court injunction that allowed them to reopen,[229] the industry's fears did not go away. Despite Mayor McClellan's failure to strong-arm the theaters in applying for more restrictive licenses, the trade association of exhibitors remained concerned.[230] The threat of regulation in New York, the most important market, was credible because, the year before, Chicago had created the first censorship board.[231]

In this context, the groups proposing voluntary measures to review the content of films started to look like attractive allies for the theater owners. Indeed, during the 1908 hearing preceding the order to close the theaters, it had become clear that the critics were divided in two camps: advocates for government censorship and those who emphasized the need for alternatives to the immoral offerings of the marketplace.[232] The People's Institute, an organization in the latter camp, proposed to create a board that would review the movies.[233] This organization was concerned about the content of the films but also believed that movies could be an important tool to educate the public.[234] For instance, the People's Institute ran a theater in New York for children that had the purpose of "educating new generations to an appreciation of the drama that will make for broader minds and more desirable citizenship."[235] Eileen Bowser notes that "[p]rogressive idealism did not conflict with [companies'] ideas of how to expand the market."[236] Cleaning up the movies could bring the movies to the middle class and expand the audience.[237] As a result, in 1909, the People's Institute and the association of exhibitors set up the New York Board of Censorship of Motion Pictures, renamed the National Board of Review in 1916.[238] It was the

227. STARR, *supra* note 25, at 306.
228. *Id.*
229. BLACK, *supra* note 25, at 13.
230. *See id.*
231. *Id.*
232. STARR, *supra* note 25, at 306.
233. BLACK, *supra* note 25, at 13.
234. *See* EILEEN BOWSER, THE TRANSFORMATION OF THE CINEMA, 1907–1915, at 38 (Charles Harpole ed., 1994).
235. *Id.* at 39.
236. *Id.* at 38.
237. *Id.*
238. FELDMAN, *supra* note 26, at 9.

industry's first attempt to ward off legal censorship through self-regulation.[239]

In parallel, the film industry became heavily concentrated. The companies Edison, Biograph, Essanay, Vitagraph, Lubin, Selig, Kalem, Melies, and Pathe had formed the Motion Picture Patents Company, also known as the Edison Trust.[240] The trust controlled the majority of the patents necessary for producing films.[241] They refused to sell raw stock film to producers outside the trust.[242] Moreover, the trust did not lease films to exhibitors that showed films made by independent producers.[243] At the moment the NBR was set up, it was still unclear whether film producers would cooperate and submit their films for review.[244] The participation of the Edison Trust was essential for the NBR to succeed.

The project of the NBR aligned well with the interests of the Edison Trust. The NBR was a means of legitimation that could defuse calls for government intervention. John Collier, leader of the People's Institute and the NBR, thought that if the industry voluntarily submitted movies for review by the NBR, the criticism would stop, and the threat of government censorship would diminish.[245] Further, among the film industry's critics, the ones setting up the NBR were the friendliest to the industry. They saw themselves as protecting film production from government censorship and serving the public by acting as mediators between film producers and society's concerns. In Professor Gregory D. Black's words, the NBR's standards for self-censorship were "just good for business."[246] The Edison Trust agreed to participate.[247]

2. The Effects of the National Board of Review

The NBR's philosophy was that regulation of movies "belonged in the hands of the public."[248] According to a public statement made by the NBR in its first year of operations, the institution would be "an agent of the public."[249] The NBR explained that its work was to serve the "common man's best interests."[250] In practice, the NBR did not allow scenes that glorified crime, vice, indecency, or immoral suggestiveness.[251] Infidelity and nudity were not allowed. Scripts that challenged the military, marriage, and the church were not advisable. The political topics and contestation of

239. *Id.* at 15.
240. *See id.* at 28.
241. FELDMAN, *supra* note 26, at 28.
242. *Id.*
243. *Id.*
244. *Id.*
245. BLACK, *supra* note 25, at 14.
246. *Id.*
247. *See* FELDMAN, *supra* note 26, at 28.
248. *Id.* at 157.
249. *Id.* at 41.
250. *Id.* The NBR also stated its purpose in a press release in 1916. *Id.* at 113.
251. *See id.* at 64–65.

traditional institutions that were common in the first motion pictures were less present during the NBR years.[252] Today, the standard sounds highly moralistic, and to some, it was.[253] Nonetheless, the NBR's members saw themselves as promoting better movies that did not appeal to the easy excitement generated by sex and violence and aimed instead at a higher artistic value.[254] They believed they were pushing studios to produce films that better served the public.[255]

The NBR's effects in terms of self-censorship must be evaluated against different baselines. Compared to films during the pre–Edison Trust period, later movies were sanitized. However, compared to the standards that a federal censorship bill would likely have adopted, the NBR was protective of free expression.[256] It actively campaigned against federal and state laws alongside the industry. Illustratively, Georgia senator Hoke Smith introduced a bill to create a Federal Motion Picture Commission in 1914.[257] His goal was to create a commission with legal authority because he believed that the NBR was too soft on censorship and too close to the industry to be independent.[258] During the hearings, film producers invoked the existence of the NBR as a sufficient filter.[259] The NBR itself organized a campaign to extend its influence to show that federal legislation was unnecessary.[260] Its project was to convince cities to allow any film approved by the NBR.[261] During this period, some new laws regulated aspects of films. For example, in 1913, the Tariff Act[262] gave the Bureau of Customs authority to inspect imported films.[263] Congress also passed a law that forbade showing actors in military uniform if the film brought dishonor or reproach upon the armed forces.[264] However, there were no national review commissions or major censorship bills.

Even though the NBR might have helped prevent more censorial federal laws, it was the Edison Trust's loss of its patents that actually restored the diverse and rich environment of the early film industry.[265] In 1912, a court overturned a patent controlled by the Edison Trust.[266] The decision allowed independent producers to use standard movie cameras.[267] Shortly thereafter,

252. STARR, *supra* note 25, at 311.
253. *Id.* at 313.
254. FELDMAN, *supra* note 26, at 40.
255. *See id.* at 41 ("[T]he Board assumed the responsibility for molding this medium in what it believed was the common man's best interests.").
256. *Id.* at 68.
257. *Id.* at 68
258. *Id.* at 69.
259. *Id.* at 71.
260. *Id.* at 78.
261. *Id.* at 79.
262. Pub. L. No. 63-16, 38 Stat. 114 (1913).
263. FELDMAN, *supra* note 26, at 66–67.
264. RANDALL, *supra* note 26, at 13.
265. STARR, *supra* note 25, at 309–10.
266. *Id.* at 309.
267. *Janet Staiger, Combination and Litigation: Structures of U.S. Film Distribution, 1891–1917*, 23 CINEMA J. 41 (1984).

independents were producing more than half of domestic films, and national production accounted for 80 percent of the market.[268] In 1915, the Motion Picture Patents Company and General Film Company were found guilty of antitrust violations and ordered to dissolve, but by then, independents had already entered the market significantly.[269]

With the dissolution of the Edison Trust, the NBR also failed. The Edison Trust, which brought together companies accounting for almost all domestic production of films,[270] had a specific interest in improving public regard for the industry. In Professor Starr's words, "an economically motivated effort to monopolize the industry produced a willing partner" in the industry for NBR cooperation.[271] Independent producers did not have the same incentives to submit their movies for review.[272] With the rise of independent producers, film production, distribution, and exhibition became fragmented, and maintaining control was challenging.[273] During this period, the dispersion of film production made it difficult for any public or private entity to control films effectively.[274]

B. The Case of Hollywood and the Hays Office

Only a few years after the fall of the Edison Trust, eight film studios grew bigger and became collectively known as the Hollywood studios.[275] With the emergence of Hollywood, pressure from religious groups, other advocates, and the government to control the content of the movies resurfaced. In response, Hollywood created the Hays Office. This new institution would function as a site for religious groups to set content rules on behalf of the interests of the public. This case study is helpful for examining a self-regulatory body in a context of highly asymmetrical lobbying.

1. The Rise of Hollywood, the Hays Office, and the Lord Code

Between 1915 and the second half of the 1920s, three developments took place in parallel: first, Hollywood gained tighter control over the industry;

268. BOWSER, *supra* note 234, at 85.
269. STARR, *supra* note 25, at 309.
270. *Id.* at 307.
271. *Id.*; *see also* Daniel Czitrom, *The Politics of Performance: From Theater Licensing to Movie Censorship in Turn-of-the-Century New York*, 44 AM. Q. 525 (1992); GARTH JOWETT, FILM: THE DEMOCRATIC ART (1976).
272. FELDMAN, *supra* note 26, at 37 (detailing the independent producers' resistance to collaborate with the NBR). The NBR also faced challenges from NBR members and associations such as the Woman's Municipal League and the Society for the Prevention of Crime, who believed that the NBR was not enough to protect children's safety. Nancy Rosenbloom, *Between Reform and Regulation: The Struggle over Film Censorship in Progressive America, 1909–1922*, 1 FILM HIST. 307, 310 (1987).
273. STARR, *supra* note 25, at 315.
274. *Id.*
275. *See generally* KIA AFRA, THE HOLLYWOOD TRUST: TRADE ASSOCIATIONS AND THE RISE OF THE STUDIO SYSTEM (2016) (detailing this history).

second, pressure to regulate the content of films and stop anticompetitive practices increased; third, Catholic groups organized and designed a strategy to participate in the governance of film production.

In the late 1920s, film studios concentrated into a small number of firms. Concentration responded to several factors. Big studios bought large theater chains to exhibit their own films, leading to vertical integration of producers, distributors, and exhibitors.[276] A system known as "block booking" forced local theaters not owned by the large production studios to book films from the large studios in blocks, leaving little space for smaller rivals.[277] Professor Starr adds that studios moved to southern California because an antiunion municipal government helped lower production costs.[278] Geographical proximity allowed industry moguls to build a community.[279] In addition, states across the country were discussing censorship bills. This led the studios to create a trade association, which further helped them coordinate anticompetitive practices. The trade association was known as the Hays Office because it was led by William H. Hays.[280]

At the beginning of his term, Hays implemented self-censorship mechanisms to accommodate the preferences of state censorship boards and avoid conflicts with religious groups.[281] The Hays Office established a voluntary mechanism for submission of film projects for review.[282] Between 1924 and 1930, it rejected 125 films it considered offensive.[283] The Office also had a department that worked directly with studios to edit any materials that could enrage state boards.[284] In 1927, the Office made a "Don'ts and Be Careful" list attempting to synthesize the rules of state boards.[285] Nudity was a "don't," and anything connected to sex, violence, and crime was included in the "be careful" column.[286]

By the end of the 1920s, however, Protestant groups believed that their criticism was insufficiently influential. They started advocating for government action to stop the block-booking practices because they forced local theaters to show immoral films.[287] At the beginning of his term, President Herbert Hoover was also evaluating antitrust action.[288] Hays was more concerned with antitrust action than with censorship regulation.[289] As a result, he was reluctant to announce that a playbook coordinated among all big industry players existed. This prevented the Office from advertising its

276. STARR, *supra* note 25, at 319.
277. *Id.*
278. *Id.* at 316.
279. *Id.*
280. *Id.* at 318.
281. *Id.*
282. *Id.*
283. *Id.*
284. *Id.*
285. *Id.*
286. BLACK, *supra* note 25, at 39.
287. STARR, *supra* note 25, at 320.
288. *Id.*
289. *Id.*

successes in self-censoring movies. In turn, the Office was unable to defend the industry from renewed accusations from religious groups of immoral content in films, concerns that were now leading these groups to lobby for the banning of block-booking practices.[290]

In this context, Catholic groups organized. Previously, the Catholic Church had not been an important actor in the space.[291] Led by Martin Quigley, this group saw an opportunity to reform the messages of films. Given the threat of antitrust action and federal censorship, Quigley envisioned a proposal to offer the film industry a self-regulatory alternative.[292] Stricter self-regulation, Quigley claimed, would benefit the industry in a number of ways. If Protestant groups were less concerned with the content of the movies, they would stop advocating for the elimination of block booking, which was the industry's top concern.[293] Quigley and Father Daniel A. Lord, a Catholic leader, would write what became the Motion Picture Production Code (the "Lord Code").[294] The Lord Code embraced a strict vision of "correct entertainment." Father Lord and Quigley shared the idea that movies should not only refrain from showing nudity, infidelity, or political issues. Movies, because of the important communications role that they had, had a "special [m]oral [r]esponsibility" to show that family, religion, and the law were the pillars of a good life.[295]

The Hays Office welcomed the Lord Code.[296] In 1930, the Hollywood studios agreed to obtain permission for new films from the Hays Office prior to production.[297] The government's tolerance of Hollywood's oligopoly was necessary to make the Hays Office's control of the industry possible.[298] According to Professor Starr, "if the industry had still consisted of dozens of production companies and distributors and hundreds of independent exhibitors, a trade association that adopted a censorship code would almost certainly have found it difficult, if not impossible, to assure compliance."[299]

2. The Effects of the Lord Code: Catholic Groups Govern on Behalf of the Public

Ultimately, the Catholic groups won. After 1934, self-censorship was very strong. According to Professor Starr, the controversial issues about class, social problems, and immigration that movies had discussed before World War I disappeared from the screen.[300] Professor Lary May has conducted an

290. *Id.*
291. BLACK, *supra* note 25, at 34.
292. STARR, *supra* note 25, at 321.
293. BLACK, *supra* note 25, at 36.
294. Francis Couvares, *Hollywood, Main Street, and the Church: Trying to Censor the Movies Before the Production Code*, 44 AM. Q. 584 (1992).
295. BLACK, *supra* note 25, at 39 (emphasis omitted).
296. STARR, *supra* note 25, at 321.
297. *Id.* at 322.
298. *See id.* at 326.
299. *Id.* at 325
300. *Id.* at 319.

analysis of the 6,600 films produced during the twenties and highlights that only fifteen (0.22 percent) dealt with labor issues, fifty-one (0.77 percent) with politics, sixty-six (0.95 percent) with religion, and 198 (2.94 percent) with issues of class or immigration.[301] Indeed, the Hays Office was explicit in its goal of keeping the industry away from controversy, not because its staffers were prudish but because their job was to avoid conflict with state governments and religious groups.[302]

One could argue that the change in the content of the films was due primarily to economic factors. Producing movies became costlier as they became longer and Hollywood stars arose. This could explain the phenomenon of programming to the middle—that is, producing uncontroversial films that would satisfy a large audience. For example, the cost of making a film in 1913 was between $1,000 and $10,000. This cost rose to between $100,000 and $500,000 in the 1920s. In 1912, sixty firms were making over 2,000 movies. In the twenties, the "Big Eight" produced 90 percent of the 800 films made yearly.[303] As a result, a small number of studios were producing most films, and they were interested in making them attractive to a wide audience.[304]

However, in 1934, the few times that the church declared a movie indecent, attendance was record-breaking.[305] Thus, the sanitization of films did not in fact respond to what most audiences would like to watch. The focus on maintaining good relationships provides a more complete explanation of sanitization in film production. Dealing with prudish activists was simply too expensive. This explanation also takes concentration as a condition of the change in movie content but emphasizes not the costs of producing movies but the costs of upsetting well-organized groups and government actors.

Comparing the screen to the publishing industry makes apparent how a small group of interests controlled film production. The courts offered protection to books to an extent not matched for movies.[306] Additionally, however, the book industry was decentralized and had no central point of control comparable to the Hays Office.[307] Catholic pressure to control movies did not exist before central control of the industry.[308] Rather, it emerged as a response to the opportunity that some leaders saw in the industry's vulnerability.[309]

301. MAY, *supra* note 224, at 214.
302. *See* BLACK, *supra* note 25, at 53.
303. MAY, *supra* note 224, at 177.
304. *Id.*
305. STARR, *supra* note 25, at 324.
306. *Id.* at 326.
307. *Id.*
308. *Id.*
309. *Id.*

C. The Case of the National Broadcasting Company and the NBC Advisory Council

The Hays Office case illustrates how a well-organized minority managed to govern the content of films, whereas the NBC Advisory Council is an example of a diverse and inclusive body representing majoritarian interests. The National Broadcasting Company set up its Advisory Council following the advice of a public relations firm hired to manage the risk of antitrust action. Its members were prestigious representatives of the major political parties, religious denominations, and other groups interested in radio programming. The Council reasoned that the public interest required airing only programs relevant to a majority of the audience. In practice, this meant that the members of the Council divided radio airtime among the powerful constituencies they represented and pushed out dissent. The next sections follow the creation of the NBC Advisory Council and its substantive outcomes.

1. The Creation of the NBC Advisory Council

The NBC emerged as a powerful actor thanks to the intervention of the Navy. After World War I, the Navy was concerned that the Great Britain-based Marconi Company would buy patent rights from America's General Electric (GE) and consolidate a worldwide monopoly over radio. The Navy wanted GE to develop a U.S. communication company. Following the Navy's advice, Owen D. Young, an attorney at General Electric, looked into the possibility of bringing U.S. firms together without violating the Sherman Antitrust Act.[310] Young was concerned about antitrust actions, but in a meeting with Attorney General Alexander Mitchell Palmer, he received informal approval for the agreement.[311] The result of these negotiations was the creation of a new company, Radio Corporation of America (RCA).

The creation of the NBC Advisory Council was the radio industry's response to the concerns raised by the press and the Federal Radio Commission over the increasingly dominant position of RCA—the parent company of NBC. Indeed, in 1926, newspapers charged that NBC would acquire a "monopoly of the air."[312] David Sarnoff, a commercial manager at RCA, hired Isroy Norr's public relations firm.[313] The firm recommended that the network create an advisory board to deflect concerns about a commercial entity exercising outsized power over public communications.[314] RCA moved quickly to create the advisory council. A letter sent on August 10, 1926, to the RCA board of directors, proposed that the advisory board function as "a court of appeal on matters of policy on programs and entertainment, as distinguished from business administration, and in cases of

310. 15 U.S.C. §§ 1–7; BENJAMIN, *supra* note 26, at 3.
311. *Id.* at 4.
312. *See, e.g.*, *Radio Corp. Takes over WEAF Station*, ROME DAILY SENTINEL, July 22, 1926, at 1.
313. BENJAMIN, *supra* note 26, at 12.
314. *Id.*

alleged discrimination."[315] In the words of Louise Benjamin, who has written extensively about the history of the Advisory Council, "the council owed much of its origin to the likelihood that charges of monopoly would be leveled at NBC and its parent company, RCA."[316] Benjamin considers that among the purposes of the creation of the Council, "deflecting cries of monopoly was paramount."[317]

The promise of the Council was to advance the public interest. At its first meeting, Young explained to the Council's founding members that the Council would advance the public interest in two ways. First, it would decide on controversies over what programs to broadcast: "[We] hope the Advisory Council may be considered as a court of appeal for complaints . . . . I should expect few will ever come to your attention unless they are really serious and difficult cases. In that case, they should be decided in the public interest."[318] Second, the Council would make recommendations to set standards for how to use airtime. Young told the council's members: "The wise guidance of able men of diversified experience located in different parts of the country is sought in order that the facilities of the National Broadcasting Company may be put to their best possible use in the public interest . . . ."[319]

Similarly to the selection of the members of the Meta Oversight Board, three considerations went into the selection of the Council's members: prestige, diversity, and collegiality. Norr recommended that the Council "have the cooperation of distinguished leaders in American public life."[320] He believed that the Council would gain legitimacy "if they selected well-recognized people who were influential, prominent, and respected within society."[321] The biographies of the members were impressive. Charles Evans Hughes had served as a Supreme Court justice and as secretary of state from 1921 to 1924.[322] John William Davis was a former U.S. solicitor general.[323] The third jurist was Elihu Root, who had served as President Theodore Roosevelt's secretary of state and later as a senator.[324]

The axes of diversity that Young considered relevant in selecting the Council's members were politics, religion, geography, and professional sector.[325] Davis was a Democrat; Root and Hughes were Republicans.[326] Julius Rosenwald, president of Sears, was to represent Jews.[327] Morgan O'Brien, a New York attorney, represented Catholics.[328] Reverend Charles

315. *Id.* at 14.
316. *Id.* at 114.
317. *Id.* at 13.
318. *Id.* at 24.
319. *Id.* at 23.
320. *Id.* at 13.
321. *Id.*
322. *Id.* at 16.
323. *Id.*
324. *Id.* at 17.
325. *Id.* at 15, 17, 19.
326. *Id.* at 16–17.
327. *Id.* at 18.
328. *Id.* at 17.

MacFarland, of the Churches of Christ in America, was to advocate for Protestants.[329] The members were selected from across the nation. Other members represented important educational and cultural institutions. Francis Farrell had been the dean of the agriculture department at Kansas State University.[330] Walter Damrosch was a conductor of the New York Symphony Orchestra, and Mary King Sherman was the president of the General Federation of Women's Clubs.[331] In addition to its determinations on religious and political controversies, the most important decisions the Council would make in the upcoming years would refer to gender and sexuality, music, and farming issues.

Finally, the Council's members were to represent diverse viewpoints but not to be dissenting voices incapable of reaching consensus with the other members.[332] In Benjamin's words, the Council's members "had to relate well with the world around them,"[333] and their views should not be "exceedingly out of step with contemporary societal or business norms."[334]

The appointment of its members anticipated the vision of the public interest that the Council would implement. Its goal was to satisfy the preferences of dominant groups equally while taming dissenting voices that could alter the carefully struck balance among powerful constituencies.

2. The Public Interest Under the NBC Advisory Council

Between 1927 and 1939, the Council had two main functions. On the one hand, it helped develop programs to advance the public interest.[335] These were called sustaining programs because they did not depend on being profitable.[336] The development of public-interest programming was fundamental for NBC's argument that reserving stations for nonprofit groups was unnecessary.[337] On the other hand, the Council's main focus was on who received access to broadcasting. The main areas where controversies arose were religious programming, partisan use of airtime, and coverage of reproductive health discussions. In these three areas, the philosophy was the same: access was distributed equally among dominant groups with different viewpoints; topics that were deemed controversial or offensive or that could upset one of the dominant groups were excluded.

329. *Id.* at 17–18.
330. *Id.* at 19.
331. *Id.*
332. *Id.* at 14.
333. *Id.* at 15.
334. *Id.*
335. *Id.* at 26.
336. *Id.* at 29.
337. STARR, *supra* note 25, at 322.

a. Religious Programming

One of the first actions of the Council was to develop principles for religious programming.[338] According to a survey that the Federal Radio Commission conducted in 1928, the average station was on the air fifty-five hours per week, eight of which were devoted to religious content.[339] The Council established principles for religious programming. NBC would serve "only the central or national agencies of the three religious faiths," religious broadcasts should "aid in popularizing religion and the church," they should "interpret religion at its highest and best," and should be done "by the recognized outstanding leaders of the several faiths."[340]

As reflected in the principles, the Council's approach to religious programming had two elements. On the one hand, it divided airtime equally among the three represented faiths (Judaism, Catholicism, and Protestantism). On the other hand, it excluded smaller groups and dissident voices. The public interest was equated with "the interest of a large number of people." Accordingly, when controversial speakers requested access to airtime, their views were, almost by definition, not in the public interest, as they did not represent the position of a large enough group.

The first controversy was brought by Joseph Franklin Rutherford of the People's Pulpit Association.[341] He asked the Council to review NBC's decision to reject his request to broadcast a religious program.[342] Briefing the Council on behalf of NBC, NBC president Merlin Hall Aylesworth explained that NBC had granted time to Rutherford following their commitment not to discriminate among the different faiths.[343] However, the speech he had prepared was a "violent attack upon organized Christianity and its churches."[344] Accordingly, NBC rejected his subsequent requests for airtime.[345]

Rutherford's message violated the Council's principles. It did not present religion at its highest, nor was it approved by the mainstream leaders of one of the dominant faiths. The Council's principles were designed to present only mainstream beliefs. Hughes noted that "certain levels of decency must be maintained" and that "NBC was 'simply discharging their obligation to the public in insisting upon the observance of that standard.'"[346] Discussing a similar issue in 1932, a member of the Council stated that stations need not be a medium for broadcasts that were "offensive to an overwhelming majority of its audience."[347]

338. BENJAMIN, *supra* note 26, at 49.
339. *Id.*
340. *Id.* at 50.
341. *Id.* at 51.
342. *Id.*
343. *Id.*
344. *Id.*
345. *Id.*
346. *Id.* at 52.
347. *Id.* at 53.

As Benjamin says, "broadcasters wanted innocuous fare, programs considered unobjectionable and safe."[348] For NBC, this approach was highly beneficial. It met the preferences of powerful groups that could protest its decisions. The Council took the same approach to partisan speech during presidential elections.[349]

b. Birth Control Coverage

The NBC Advisory Council relied on its definition of the public interest as the interests of the majority to deny airtime to women advocates discussing birth control. In the late 1920s, the American Birth Control League (ABCL), which would become the Planned Parenthood Federation of America ("Planned Parenthood") in 1942, sought access to broadcasting to cover debates about birth control.[350] Newspapers covered the controversies widely; however, securing access to broadcasting was difficult. NBC rejected ABCL's request to cover its 1929 three-day convention on birth control.[351] As in other areas, most broadcasters followed NBC's approach.[352] The ACLU appealed the decision to the NBC Advisory Council and to the Federal Radio Commission on behalf of the ABCL.[353] The Radio Act of 1927[354] had mandated that radio stations act in the public's interest.[355] The issue was whether coverage of birth control was in the public interest.[356]

The Council decided the controversy in favor of NBC. The ACLU's position was that radio stations were licensees using a public resource and, accordingly, had to broadcast all viewpoints on controversial subjects of public importance.[357] The ACLU's complaint stated that NBC "neglected the performance of a public service which it was under a sort of moral obligation to perform" and that it was improper for NBC to censor material.[358]

The Council characterized itself as a "public trustee."[359] It considered that its role in this capacity was to decide whether a majority of listeners would welcome the coverage of birth control. Hughes argued that NBC's duty to the public was to respect the audience's desires and avoid programs that some might not have received well.[360] Explaining why NBC had made the right decision, Young explained, "I feel sure that a referendum vote of listeners

348. *Id.* at 64.
349. *Id.* at 78.
350. *See, e.g.*, CAROLE R. MCCANN, BIRTH CONTROL POLITICS IN THE UNITED STATES, 1916–1945, at 15–17 (1994); BENJAMIN, *supra* note 26, at 67.
351. BENJAMIN, *supra* note 26, at 68.
352. *Id.* at 66, 74.
353. *Id.* at 68.
354. Pub. L. No. 69-632, 44 Stat. 1162 (1927).
355. *Id.* § 4.
356. BENJAMIN, *supra* note 26, at 68.
357. *Id.* at 69.
358. *Id.* (referencing a letter from the ACLU to John Elwood on November 22, 1929).
359. *Id.* at 70.
360. *See id.* at 73 (citing "Detailed Report" from 1930).

would have demonstrated overwhelmingly that it was the kind of subject which was not yet ripe for introduction through the radio to the homes of America."[361] Young considered that the movement had "not progressed far enough in public interest and good-will to warrant its broadcast by such a trustee of wave-lengths as the National Broadcasting Company."[362]

c. Sustaining Programs

The NBC Advisory Council designed two programs to serve the public interest: the "Music Appreciation Hour" and the "National Farm and Home Hour."[363] They did not depend on advertising revenue and were instead supported by NBC. Their goal was to serve rural audiences.[364] For example, the National Farm and Home Hour contained a mix of weather reports, debates on legislation affecting the agricultural sector, and entertainment.[365] The federal government collaborated with NBC in developing the programming, and the collaboration was beneficial for both parties. The U.S. Department of Agriculture (USDA) provided informational material used for the information segments of the show.[366] The U.S. Weather Bureau was in charge of a daily weather forecast.[367] The relationship gave the government some control over content at a time when farm issues were political and part of the 1932 presidential campaign. In addition, other governmental bodies had access to time on the program. For example, the Post Office presented weekly updates, and the Department of the Interior ran a series of episodes on medical care in rural areas.[368]

The biggest payoff for NBC came in 1934 when the administration of Franklin D. Roosevelt was considering a reform in communications policy.[369] Josephus Daniels, a close advisor, had suggested that President Roosevelt consider nationalizing radio.[370] The President rejected the proposal but asked Commerce Secretary Daniel C. Roper to assess the communication infrastructure.[371] The resulting report recommended the creation of a regulatory agency for communications policy.[372] During the debate about this new commission, supporters of reform proposed reserving 25 percent of channels for nonprofit broadcasters.[373] Ultimately, Congress established the Federal Communications Commission in the Communications Act of 1934.[374] The bill included nothing to modify

361. *Id.* at 70.
362. *Id.* (referencing a letter from Owen Young to Harry Ward on December 3, 1929).
363. *Id.* at 26–47.
364. *Id.* at 26.
365. *Id.* at 28.
366. *Id.* at 33.
367. *Id.*
368. *See id.* at 35.
369. *See* STARR, *supra* note 25, at 359–60.
370. *Id.*
371. *Id.* at 360.
372. *Id.*
373. *Id.*
374. Pub. L. No. 73-416, § 1, 48 Stat. 1064, 1064 (1934).

broadcasting infrastructure.[375] The USDA testified in favor of maintaining the status quo. It argued that the "Farm Hour" and "Music Hour" showed that NBC could support public-interest programming within the current setup.[376]

According to Professor Starr's analysis, President Roosevelt had two additional reasons not to intervene in the structure of broadcasting. First, it was not a policy priority. Second, NBC had been receptive to demands from the government for airtime and never showed any intention of jeopardizing official communications.[377] As Professor Starr concludes, "[u]nlike other areas of policy, broadcasting would have no new deal."[378]

In all these cases, self-regulation was the companies' response to their institutional vulnerability. This might seem counterintuitive since these companies were first and foremost greatly powerful. However, the prospect of losing a beneficial legal environment made them effective targets for lobbying pressure. Under these conditions, advocates had the chance of gaining a foothold in speech governance.

The effects of these partnerships between companies and advocates are disappointing. In highly asymmetrical contexts such as the religious organizing around Hollywood studios, the self-regulatory bodies' claim of governing on behalf of the public interest was never credible. Of course, depending on which group has asymmetric power, the substantive outcomes might be better or worse. The point is that the process itself is epistemically unsound.

Even the NBC Advisory Council, with its arguably satisfactory diversity,[379] failed to serve as an effective counterweight to corporate power capable of protecting dissent and creating avenues for minorities to contest the hegemonic order of the public sphere. In fact, the NBC Advisory Council and the Meta Oversight Board illustrate the impossibility of consensus without exclusion.[380] In the case of the NBC Advisory Council, that is evident as the Council's vision of the public interest was explicitly to avoid controversy.[381] The cost of having representatives of all major viewpoints come to an agreement was to leave out unsettling viewpoints.[382] The Meta Oversight Board is more sophisticated, as it has adopted a framework—IHRL—that protects dissent strongly.[383] Nevertheless, the Board is tolerant of nudity bans and other prohibitions on speech that international law protects.[384] More fundamentally, its structure is not designed to facilitate

375. STARR, *supra* note 25, at 361.
376. *See* BENJAMIN, *supra* note 26, at 39.
377. *See* STARR, *supra* note 25, at 360.
378. *Id.* at 361.
379. At least considering the criteria considered relevant at the time. The members of the Council were not diverse in terms of race and gender.
380. *See* MOUFFE, *supra* note 22, at 92.
381. *See supra* Part III.C.2.
382. *See supra* Part III.C.2.b.
383. Strossen, *supra* note 74, at 331.
384. *See Breast Cancer Symptoms and Nudity*, *supra* note 135; *Depiction of Zwarte Piet*, *supra* note 181; *South Africa Slurs*, *supra* note 90.

power-sharing but to attain authority through claims of objective decision-making. Bodies of responsible, well-intentioned, and diverse advocates that aim at performing as objective decision-makers are ill positioned to protect dissent and to participate themselves as political actors in the governance of the public sphere. As a result, the protection of contestation must come from the legal architecture of the public sphere.

On a more hopeful note, the case studies not only illustrate the problematic effects of the inclusion of advocates in speech governance but also point in the direction of an alternative path. In each case, there was a discussion about a relevant legal intervention that had the potential to disrupt corporate power and support the construction of a different public sphere. A legal challenge to the Edison Trust's patent pool brought diversity back to the screen.[385] During the Hays Office period, the proposals to ban block-booking practices had the potential to disrupt the power of Hollywood studios and the effectiveness of the partnership between the Hays Office and Catholic groups.[386] NBC invoked the work of its Advisory Council to stall legal proposals to redistribute the radio spectrum and support the creation of nonprofit radio stations.[387] These stations could have been assigned to dissenting groups, some of which might have been willing to cover the conferences on birth control, as the ACLU and Planned Parenthood wanted.[388]

Taking inspiration from the case studies, the next part examines how to use law to construct a public sphere that more forcefully protects the contestation of its own hegemonic order.

IV. TOWARD PLURALISM IN THE DIGITAL PUBLIC SPHERE

As a response to the exclusionary effects of the new governance project for content moderation, this part explores pluralism as an alternative vision for speech governance. Given that self-regulation and multistakeholder governance cannot live up to their promises, this part considers how the law might create alternative participatory structures.

Pluralism advances the commitment to participatory parity more meaningfully for traditionally excluded constituencies. A plural public sphere does not expect excluded groups to gain inclusion in a negotiated process to achieve reconciliation with more powerful actors.[389] Rather, it gives these groups a space that is simultaneously separate from and

385. *See supra* Part III.A.2.
386. *See supra* Part III.B.2.
387. *See supra* Part III.C.2.c.
388. *See supra* Part III.C.2.b.
389. *See* MOUFFE, *supra* note 22, at 92 (describing an agonistic public sphere as a space "where conflicting points of view are confronted without any possibility of a final reconciliation"); *see also* Scott Skinner-Thompson, *Agonistic Privacy & Equitable Democracy*, 131 YALE L.J.F. 454, 456 (2021).

connected to mainstream spaces, so they can function as sites both for regroupment and contestation.[390]

Part IV.A argues that a plural public sphere can advance the ideal of participatory parity in a more promising way than the new governance model for content moderation. Part IV.B suggests what a plural platform-based public sphere might look like in practice. Part IV.C examines the legal techniques necessary for implementing pluralism today.

A. In Defense of Pluralism in the Digital Public Sphere

For dissenting groups, pluralism can sometimes offer a more empowering tool to participate in the production of social meaning and contest the hegemonic order of the public sphere than the new governance model for content moderation. Although some advocates might prefer reaching compromises with the dominant actors, for those who are less likely to persuade companies or more powerful advocates, building mechanisms outside of a negotiated process holds more promise than trying to gain inclusion in the governance elite.[391]

Pluralism favors excluded groups in two key ways. First, and more obviously, it awards them a space of their own control. Instead of forcing them to enter negotiations with groups they cannot persuade or to reshape their interests as to reach compromises, they can adopt their own rules and prioritize their own interests. Excluded groups can use these spaces to produce counter narratives and interpretations of their discursive realities. Second, in a decentralized environment, it becomes challenging for governments and advocacy groups to exert lobbying pressure for the market to reflect their preferences and exclude unwanted speakers.

On this last point, industries with clear checkpoints are institutionally more vulnerable to pressure to exclude controversial speech. Historically, more competitive spaces have diffused the power of specific groups to extend their influence. Remember the comparison between the effects of Catholic and Protestant lobbying efforts to censor immoral stories in the movie and publishing industries.[392] The creation of the Hays Office allowed religious groups to articulate and target their demands. In turn, the Hays Office accepted to incorporate those demands as rules for producing movies believing that, in exchange, religious groups would stop calling for antitrust enforcement actions. The book industry did not have a comparable central entity and did not become an intense target of religious pressure during that time.[393] Similarly, while the NBC Advisory Council refused to broadcast

390. *See* Fraser, *supra* note 22, at 68.

391. Fraser, *supra* note 22, at 66 ("I contend that, in stratified societies, arrangements that accommodate contestation among a plurality of competing publics better promote the ideal of participatory parity than does a single, comprehensive, overarching public.").

392. *See supra* Part III.B.2.

393. *See* STARR, *supra* note 25, at 326.

discussions about reproductive health as requested by Planned Parenthood and the ACLU, newspapers covered the topic widely.[394]

Pluralism is beneficial when it creates simultaneously separate and interconnected spaces.[395] Indeed, pluralism as the mere existence of various publics is ever present. Dissenting, subordinate, or counter publics have historically set up spaces for talk and, contemporarily, it is very much possible for any community to create an online space of their own.[396] Yet, as Professor Nancy Fraser articulates, the emancipatory potential of pluralism resides in spaces that can serve as "spaces of withdrawal" and "bases . . . for agitational activities directed toward wider publics."[397] This point is key. The challenge is to create spaces for regroupment that can still serve as the basis for contestation of dominant discourses. Thus, the task is to create plurality within mainstream platforms or forms of connecting contesting spaces to larger audiences.

Pluralism posits difficult challenges. Some might worry that more diversity could worsen fragmentation, siloing, and echo chambers.[398] This fear stems from one of the implicit premises of the new governance project for content moderation: that a unified and comprehensive public sphere is superior to a multiplicity of spaces in conflictual relations. That is precisely the fundamental premise that this Article challenges.[399] Relatedly, pluralism means making space for speakers that are not necessarily virtuous and might even be against equality, freedom of speech, or political participation. Admittedly, pluralism can broaden discursive contestation in all directions. Yet, making space for groups can be deradicalizing. Professor Chantal Mouffe argues that creating space for conflict—even when making space for far-right voices—keeps those voices engaged in mainstream conversations and ultimately committed to the project of collective governance.[400] Beyond these theoretical defenses of pluralism, the suggestions for implementation in the next section also offer a practical response to these concerns. The next section advocates for a reconfiguration of social media that would alter the

394. *See supra* Part III.C.2.b.

395. *See* Fraser, *supra* note 22, at 68 (identifying these two functions of what Professor Nancy Fraser calls "subaltern counterpublics").

396. *See id.* at 61.

397. *Id.* at 68 (identifying an analogous premise in Habermas' conceptualization of the ideal public sphere).

398. *See generally* CASS SUNSTEIN, #REPUBLIC: DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA (2018) (articulating the concern about fragmentation).

399. *See also* Amin Mahmoudi, Dariusz Jemielniak & Leon Ciechanowski, *Echo Chambers in Online Social Networks: A Systematic Literature Review*, 12 IEEE ACCESS 9594, 9616 (2024) (showing that the effects of online echo chambers on public discourse remains controversial).

400. *See* CHANTAL MOUFFE, FOR A LEFT POPULISM 22 (2018); *see also* Shiza Ali, Mohammad Hammas Saeed, Esraa Aldreabi, Jeremy Blackburn, Emiliano De Cristofaro, Savvas Zannettou & Gianluca Stringhini, *Understanding the Effect of Deplatforming on Social Networks*, 12TH ACM WEB SCI. CONF. 187, 192 (2021) (showing that when users are banned from Reddit and Twitter and migrate to smaller platforms, they tend to become more toxic).

algorithmic patterns of speech distribution and, thus, the infrastructural roots of misinformation, radicalization, polarization, and other social harms.[401]

B. What Pluralism Could Look Like: Competition Among a Plurality of Recommendation Systems

How to recreate pluralism on social media is not at all obvious. One particularly exciting project is to introduce an editorial layer between platforms and users. This approach envisions a separation of content moderation from content hosting services, where existing platforms would solely host content generated by users without organizing its distribution. On top of platforms, multiple entities would offer competing recommendation systems prioritizing diverse interests and preferences.

A number of prominent commentators, such as Daphne Keller and Francis Fukuyama, have anticipated this approach.[402] In this reordered public sphere, users would select a recommendation system to display content upon entering their preferred social media platform, akin to choosing a channel on a television. Although a platform like Facebook would continue to host content generated by users, various corporate, nonprofit, and governmental actors would provide multiple options to organize users' feeds. For example, imagine Planned Parenthood had its own recommendation system featuring information about reproductive health and content generated by prominent pro-choice advocates. Other options might prioritize specific themes like right-wing content, political content more generally, sports, or other forms of entertainment. Interested users could select Planned Parenthood's channel sometimes and easily switch to other channels when looking for something else. Indeed, as shown in the experiences of the film and radio, competition often results in speech diversity.[403] A robust project should include public or state-supported alternatives to ensure that groups with fewer resources can control parts of the digital public sphere. This would provide previously excluded groups with an opportunity to build spaces that reflect their values and prioritize their own content.

Like any proposal to reimagine social media platforms, this project entails many risks: concentration might reemerge, competition might be insufficient to include marginalized groups, and giving access to new actors to social media infrastructure might amplify privacy and security concerns. These are serious concerns that might make the project ultimately futile. Yet, a lot can be done to address them. I discuss some of these challenges and possible remedies in the last section. Here, I am interested in how pluralism positively

401. *See* JOSH SIMONS & DIPAYAN GHOSH, BROOKINGS INST., UTILITIES FOR DEMOCRACY: WHY AND HOW THE ALGORITHMIC INFRASTRUCTURE OF FACEBOOK AND GOOGLE MUST BE REGULATED (2020), https://www.brookings.edu/wp-content/uploads/2020/08/FP_20200908_facebook_google_algorithm_simons_ghosh.pdf [https://perma.cc/4J7W-EQ48]; *see also* Cohen, *supra* note 120, at 4.

402. *See supra* note 33.

403. *See generally* EDWIN BAKER, MEDIA, MARKETS, AND DEMOCRACY (2002) (examining how market structure relates to diversity in programming).

affects the relationship between firms and advocates—which gives reason to find the project attractive.

In this new institutional context, advisory bodies would not disappear. These entities might help companies offering recommendation systems build trust, attract specific audiences, or might still be the response to public demands for responsible corporate behavior. In a plural environment, these bodies would play a more productive role. They would not be deployed to prove that companies can figure out what the public wants without the help of the state. Instead, they could be explicit about their normative values. The legitimacy of their decisions would depend not on proving that they can act on behalf of everyone but on doing a good job at representing *someone*. Thus, advocates might be able to make more committed decisions in difficult cases similar to the Meta Oversight Board's decision regarding the deplatforming of Trump. As discussed above, in that opportunity, the Board omitted taking any forceful stance to avoid being perceived as a political actor. In a competitive and plural environment, advocates would be expected to explain why their approach is superior to others.

Thus, introducing more diversity in the online public sphere should not be seen as undermining the role of bodies like the Meta Oversight Board but as creating the right institutional context for these initiatives to meet their most important goal—namely, to provide public explanations for how they decided to balance competing values.[404] Their decisions would be, in turn, subjected to contestation by other actors. Ultimately, advisory bodies would be participating in a space designed for confrontation and persuasion among well-reasoned viewpoints.

In this world, many of the tools of the new governance project, far from becoming irrelevant, would play a vital role. Transparency and audits would enable the public and the state to understand and regulate the media environment.[405] Disclosure of the contacts between state actors and companies—especially cases where state actors request the removal of content—would facilitate public oversight of state influence over the public sphere.[406]

Finally, it is worth noticing that introducing competition at the level of recommendation systems is more promising than other existing proposals to encourage diversity. For example, distributed platforms gained popularity when Elon Musk took over Twitter.[407] As a form of resistance, users flocked to Mastodon, a platform that presents itself as resilient against centralized power.[408] Decentralized or distributed social media do not have a central authority that can remove content from the entire network or exclude any

404. *See* Feldman, *supra* note 3, at 104.

405. *See* Lobel, *supra* note 9, at 345.

406. *See* Douek, *supra* note 3, at 591.

407. Alan Z. Rozenshtein, *Moderating the Fediverse: Content Moderation on Distributed Social Media*, 3 J. FREE SPEECH 217, 218 (2023).

408. Barbara Ortutay, *Twitter Drama Too Much?: Mastodon, Others Emerge as Options*, ASSOCIATED PRESS (Nov. 12, 2022), https://apnews.com/article/elon-musk-twitter-inc-business-6e8bff97bb33d53b87c3d9a7c6a509cf [https://perma.cc/97HK-WCYF].

person from participating in it.[409] After the initial enthusiasm following Elon Musk's purchase of Twitter, the number of active users on Mastodon has dropped.[410] Thus far, uptake has been a challenge.[411] Some of the uncritical enthusiasm around distributed platforms resembles the truncated hopes around the early internet era.[412] Left to the market's doing, decentralized technology did not stand in the way of the development of new forms of concentrated private power.[413]

The project to break up platforms gained momentum in 2020, when Senator Elizabeth Warren was campaigning on breaking up big tech.[414] What began as a radical project quickly diluted into a very modest idea. The core of the proposal became separating Facebook from Instagram and WhatsApp, which would not meaningfully address concerns about concentration.[415] Even if it was plausible to break these platforms into smaller pieces, users have shown strong preferences for spaces where most people can gather rather than for fragmented and disconnected platforms. Unlike the rudimentary idea of breaking up platforms into small fragments, algorithmic diversity can function on large platforms where all content exists.

C. Using Law to Reorder the Public Sphere: The Turn Toward Infrastructure Regulation

Law offers useful tools to change the configuration of the public sphere and encourage the development of pluralism. In order to make pluralism a reality, Congress could mandate a structural separation between hosting and distribution services on social media. Platforms would be dedicated to host content generated by users, but not to organize its distribution. The law should categorize them as common carriers, that is, should prohibit them from making decisions about what speech to allow or disallow. However, there could be certain exceptions to that principle for specific categories of

409. Chand Rajendra-Nicolucci & Ethan Zuckerman, *What If Social Media Worked More Like Email?*, *in* AN ILLUSTRATED FIELD GUIDE TO SOCIAL MEDIA 24, 24–25 (Chand Rajendra-Nicolucci & Ethan Zuckerman eds., 2021).

410. Amanda Hoover, *The Mastodon Bump Is Now a Slump*, WIRED (Feb. 7, 2023), https://www.wired.com/story/the-mastodon-bump-is-now-a-slump/ [https://perma.cc/G8MT-8RP5].

411. *See* Alan Rozenshtein, *Mastodon's Content-Moderation Growing Pains*, VOLOKH CONSPIRACY (Nov. 21, 2022), https://reason.com/volokh/2022/11/21/mastodons-content-moderation-growing-pains/ [https://perma.cc/83W4-FZLE].

412. *See* Yochai Benkler, *The Role of Technology in Political Economy: Part 2*, LAW & POL. ECON. BLOG (July 26, 2018), https://lpeproject.org/blog/the-role-of-technology-in-political-economy-part-2/ [https://perma.cc/3N65-AJNY].

413. *See* Amy Kapczynski, *The Law of Informational Capitalism*, 129 YALE L.J. 1460, 1493 (2020).

414. Elizabeth Warren, *Break Up Big Tech*, WARREN, https://2020.elizabethwarren.com/toolkit/break-up-big-tech [https://perma.cc/56U4-REB7] (last visited Feb. 14, 2025); Sheelah Kolhatkar, *How Elizabeth Warren Came Up with a Plan to Break Up Big Tech*, NEW YORKER (Aug. 20, 2019), https://www.newyorker.com/business/currency/how-elizabeth-warren-came-up-with-a-plan-to-break-up-big-tech [https://perma.cc/6VKP-PZKK].

415. *See, e.g.*, Evelyn Douek, *Breaking Up Facebook Won't Fix Its Speech Problems*, SLATE (May 10, 2019), https://slate.com/technology/2019/05/chris-hughes-facebook-antitrust-speech.html [https://perma.cc/3XL4-PKY2].

illegal speech such as child sexual abuse material.[416] Distribution services would organize how content is prioritized and presented to users. They would need access to platforms' infrastructure in order to operate on them. The same statute would create an agency or give the Federal Communications Commission new authority to administer certain aspects of this new market of distribution services.

This proposal follows a tradition in U.S. law that protects free speech values by regulating infrastructure to promote competitive markets.[417] In a context where many projects raise First Amendment concerns, this one does not require regulating speech.[418] On the contrary, it promotes competition and the marketplace of ideas that the First Amendment envisions.[419] The Supreme Court's recent decision in *NetChoice* could suggest that any legal obligation for platforms to host content would violate the First Amendment. However, as Justice Alito points out in his concurring opinion, the American tradition includes the enactment of legal requirements to carry speech and content.[420] This tradition includes the regulation of the postal service, the telegraph, the telephone, and even railroads.

Perhaps surprisingly, the Hepburn Act of 1906,[421] which regulated the railroad industry, offers a useful example of structural separation.[422] At the turn of the twentieth century, railroad companies used to own or control most of the country's coal mines. They favored their own companies and disadvantaged independent coal competitors in the transportation of commodities.[423] In response, the Hepburn Act banned railroad companies from shipping commodities produced by them. In other words, it imposed a structural separation between companies' railroads (infrastructure) and commercial activities (content) to promote competition in the commodities' sector.

Relying on an analogous technique, the Post Office Act of 1792[424] (the "Act") regulated the postal service to promote the free diffusion of knowledge.[425] At the time, the mail was an essential information network. The Act granted all newspapers the right to access the service; it prohibited

416. *See* Ganesh Sitaraman, *Deplatforming*, 133 YALE L.J. 497, 502 (2023) ("[I]n American law, the duty to serve all comers was never absolute.").

417. *See, e.g.*, Genevieve Lakier, *The Non-First Amendment Law of Freedom of Speech*, 134 HARV. L. REV. 2299, 2371–76 (2021).

418. *See* Daphne Keller, *Amplification and Its Discontents*, KNIGHT FIRST AMEND. INST. (June 8, 2021), https://knightcolumbia.org/content/amplification-and-its-discontents [https://perma.cc/T2U8-KCJJ].

419. *See* Associated Press v. United States, 326 U.S. 1 (1945) (holding that the Sherman Antitrust Act could require the Associated Press to share its news with newspapers that competed with its members).

420. Moody v. NetChoice, LLC, 144 S. Ct. 2383, 2422 (2024) (Alito, J., concurring in the judgment).

421. Pub. L. No. 59-337, 34 Stat. 584 (1906).

422. *Id.*

423. *See* MORGAN RICKS, GANESH SITRARAMAN, SHELLEY WELTON & LEV MENAND, NETWORKS, PLATFORMS, AND UTILITIES LAW AND POLICY 487 (2022).

424. Pub. L. No. 2-7, 1 Stat. 232 (1792).

425. *See* Lakier, *supra* note 417, at 2309.

postal agents from inspecting mail (and, thus, discriminating on the basis of content); and established subsidized rates for newspapers.[426] These conditions enabled a competitive market of newspaper publishers.[427] Furthermore, by granting every newspaper a right of access to the postal service, the Act prevented postal agents who were also newspaper publishers from favoring their own publications while denying postal service to their competitors.[428]

Congress used the same technique to regulate the telegraph and the telephone. Section 202 of the Communications Act[429] prohibits these industries—and all information services defined as common carriers—from discriminating against consumers based on the content of their speech.[430] This neutrality requirement helps ensure that "many types of life, character, opinion and belief can develop unmolested and unobstructed."[431] Indeed, infrastructural regulation has been an effective tool that Congress has relied on to encourage competition and safeguard the pluralism of ideas.

Other authors have suggested that unbundling requirements could be the most suitable legal technique to introduce competition.[432] That is an understandable suggestion because unbundling requirements force incumbents to license hard-to-duplicate infrastructure to competitors. The best-known example of unbundling requirements is the Telecommunications Act of 1996.[433] It forced local telephone companies to share their infrastructure with newcomers. However, unbundling requirements failed to promote competition in the telephone industry.[434] Incumbents were forced to share their networks with direct competitors and, rationally, made significant efforts to block that access. A structural separation approach is more promising since it would take existing platforms out of the content moderation business. Thus, incumbents would have fewer incentives to deter newcomers' access because they would not be direct competitors.[435]

In the case of social media platforms, a neutrality requirement or must-carry duties are touted on both sides of the political spectrum to protect different forms of expression on privately owned platforms.[436] However,

426. *See* Pub. L. No. 2-7, 1 Stat. 232 (1792).
427. STARR, *supra* note 25, at 251.
428. *See* Lakier, *supra* note 417, at 2310.
429. Pub. L. No. 73-416, § 202, 48 Stat. 1064, 1070 (1934).
430. 47 U.S.C. § 202; *see also* Cantwell v. Connecticut, 310 U.S. 296, 308 (1940) (interpreting that this neutrality requirement protects First Amendment interests).
431. *Cantwell*, 310 U.S. at 310.
432. Keller, *supra* note 33, at 168.
433. Pub. L. No. 104-104, § 271, 110 Stat 56, 86–92 (1996).
434. Gene Kimmelman, Mark Cooper & Magda Herrera, *The Failure of Competition Under the 1996 Telecommunications Act*, 58 FED. COMM. L.J. 511, 511 (2006); RICKS ET AL., *supra* note 423, at 389.
435. I am thankful to Morgan Ricks for raising this point.
436. *See, e.g.*, Eugene Volokh, *Treating Social Media Platforms Like Common Carriers?*, 1 J. FREE SPEECH L. 377 (2021); Genevieve Lakier & Nelson Tebbe, *After the "Great Deplatforming": Reconsidering the Shape of the First Amendment*, LPE BLOG (Mar. 1, 2021), https://lpeproject.org/blog/after-the-great-deplatforming-reconsidering-the-shape-of-the-firs

unlike the telegraph or the telephone, an obligation for platforms not to discriminate would not be enough to generate productive pluralism in this context. For platforms to be usable and valuable, it is necessary that intermediaries organize how content is displayed.[437] Thus, the neutrality of platforms must be complemented with a structure that enables those interested in offering recommendation systems to access the platform where they want to run their services and to the data that feeds recommendation systems so that they can build their own.

This new market of distribution services or recommendation systems presents two challenges that law can help address: competition and privacy concerns.

First, without adequate public oversight, recentralization of corporate power on this new market of recommendation systems is highly probable. For example, if large platforms vertically integrate with their default algorithm, change is unlikely. Similarly, if mergers and acquisitions are not properly supervised, this new market could quickly go through the same recentralization patterns observed with social media companies. In order to foster relevant plurality, it would be necessary to supervise mergers and strategies to stifle competition effectively.

Second, opening the access to people's data to multiple actors is also a challenge. Beyond adequate federal legislation of data protection,[438] Congress should create a new agency or give an existing agency new authority to control access to social media infrastructure. The Hepburn Act and the Telecommunications Acts that regulated other platforms relied on an agency to administer various aspects of the legislative framework. In this case, an administrative agency should ensure that distribution services that require access to platforms' data meet legal conditions to properly safeguard and use that data. This agency could also regulate the prices that platforms might charge to license their infrastructure in case commercial actors do not reach an agreement.[439]

Ultimately, reordering the public sphere requires a robust role for the state. The new governance paradigm in its original formulation—before it was imported to the content moderation field—attributed important functions to the state. The state was not only to support the negotiated process among stakeholders, but also to participate in helping define policy goals. Professor Freeman highlights that these engagements show that the role of the state in this model is not small—rather, through these negotiations, it can regulate

t-amendment/ [https://perma.cc/8GDT-2MUE]; Simons & Ghosh, *supra* note 401; Waldman, *supra* note 38.

437. *See* TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET 5–6 (2018).

438. *See* Julie Cohen, *How (Not) to Write a Privacy Law*, KNIGHT FIRST AMEND. INST (Mar. 23, 2021), https://knightcolumbia.org/content/how-not-to-write-a-privacy-law [https://perma.cc/V6S2-RRWM].

439. Similarly, the Telecommunications Act of 1996 gave authority to the Federal Communications Commission to set the rates charged by the incumbents to newcomers for the leased infrastructure. *See* Verizon v. FCC, 535 U.S. 467 (2002).

areas that it cannot lawfully reach directly.[440] In the context of speech governance, however, this form of involvement would be too far reaching. State participation in the coregulation of lawful speech is precisely what the new governance project for content moderation seeks to prevent in the first place.[441] From this perspective, encouraging governments to negotiate speech rules circumventing the democratic lawmaking process would be an even worse scenario.[442] Thus, turning to infrastructure regulation more adequately recenters the state in this specific context.

Despite the challenges, the turn toward infrastructure regulation holds great promise. One could be skeptical that competition could be an effective mechanism to discipline the social media market. Yet, the proposal to reintroduce diversity in the platform-based public sphere is not a project that prioritizes market competition as a means to postpone state regulation. On the contrary, it is part of a move toward looking at the social media infrastructure (including how users' data is collected, who has access to it, and how companies use it to structure the flow of information) as a means to enable self-governance through democratic institutions. Indeed, the regulation of speech infrastructures shows that democratic institutions can significantly embed values in the public sphere without engaging in unlawful speech regulation.

Focusing the attention on infrastructure is a way out of the false dilemmas at the ideological foundation of the new governance project for content moderation. Part I analyzed how the juxtaposition of multiple problems with social media, ranging from ownership concentration to the spread of harmful speech, has been unproductive. At first sight, these problems point in conflicting directions and lead to the new governance project for content moderation as the only alternative. However, tracing these problems back to the algorithmic order of the public sphere gives a coherent point of intervention that can address many of these concerns. On the one hand, using law to structure the public sphere is a promising response to the excessive accumulation of corporate power. Simultaneously, it can ameliorate the concerns about how problematic speech spreads.

Regarding concerns about misinformation and hateful speech, social media companies have successfully positioned themselves as the only actors who can control speech that is harmful but legal. Yet, recommendation algorithms are responsible for the excessive distribution of harmful speech in the first place.[443] Competition alone might not necessarily address those

440. *See* Freeman, *supra* note 8, at 671.

441. *See* Minow & Minow, *supra* note 7, at 432.

442. *See generally* Derek Bambauer, *Against Jawboning*, 100 MINN. L. REV. 51 (2015).

443. *See* Nadler, Crain & Donovan, *supra* note 114; Miriam Fernández, Alejandro Bellogín, Iván Cantador, *Analysing the Effect of Recommendation Algorithms on the Amplification of Misinformation*, ACM WEB SCI. CONF., May 21, 2024; Hussein Eslam, Juneja Prerna & Mitra Tanushree, *Measuring Misinformation in Video Search platforms: An Audit Study on YouTube*, PROC. ACM HUM.-COMPUT. INTERACT., May 2020; Jennifer Cobbe & Jatinder Singh, *Regulating Recommending: Motivations, Considerations, and Principles*, 10 EUR. J.L. & TECH 1, 7 (2019); Narayanan, *supra* note 110, at 35; Nathalie Maréchal, Tebecca MacKinnon & Jessica Dheere, *Getting to the Source of Infodemics: It's the Business Model*,

concerns.[444] Infrastructure regulation enables projects aiming at changing the algorithmic patterns of speech distribution, for example by regulating targeted advertising, how companies gather and aggregate users' data, and process it for advertising practices and algorithmic optimization.[445]

Overall, the turn toward infrastructure offers an entry point for interventions to decide collectively what values should be embedded in the production and distribution of public discourse. Structural separation between platforms and recommendation systems might be a means to encourage pluralism and contestation.

CONCLUSION

There are growing concerns about the values that guide content moderation. Many believe that speech governance should not only respond to business interests but also advance democratically defined priorities and offer strong opportunities for the contestation of the hegemonic order of the public sphere. The most promising way to achieve this is by regulating the architecture of the public sphere—reintroducing diversity in speech governance, supporting public options, and regulating how companies use people's data to distribute speech and sustain their advertising practices. However, the proposals gaining traction among lawmakers and scholars instead prioritize audited self-regulation, disclosure regimes, and the inclusion in speech governance of advocates as representatives of the public. This approach is appealing because it promises to counter corporate power without involving the state in the speech realm. However, an institutional understanding of this model reveals that aiming at defining public goals for speech governance through a negotiated practice between companies and advocates is conceptually incoherent. Historical and contemporary examples further illustrate that this model bends governance toward the positions of advocates selected by companies for their relative friendliness to corporate interests. Despite the insistence on making self-regulatory bodies more independent, diverse, or inclusive, this governance model is inherently incompatible with its own stated goal of constructing public values. The project is, thus, ineffective as a counterweight to corporate power and might

---

NEW AM. (May 27, 2020), https://www.newamerica.org/oti/reports/getting-to-the-source-of-infodemics-its-the-business-model/ [https://perma.cc/243W-6DHH]; Tomer Shadmy, *Content Traffic Regulation: A Democratic Framework for Addressing Misinformation*, 63 JURIMETRICS J. 1 (2022); Susan Benesch, *Proposals for Improved Regulation of Harmful Content Online*, *in* REDUCING ONLINE HATE SPEECH: RECOMMENDATIONS FOR SOCIAL MEDIA COMPANIES AND INTERNET INTERMEDIARIES 247 (Yuval Shany ed., 2020).

444. *See* Genevieve Lakier, *The Limits of Antimonopoly Law as a Solution to the Problems of the Platform Public Sphere*, KNIGHT FIRST AMEND. INST. (Mar. 20, 2020), https://knightcolumbia.org/content/the-limits-of-antimonopoly-as-a-solution-to-the-problems-of-the-platform-public-sphere [https://perma.cc/U77N-8P65].

445. *See* K. Sabeel Rahman & Zephyr Teachout, *From Private Bads to Public Goods: Adapting Public Utility Regulation for Informational Infrastructure*, KNIGHT FIRST AMEND. INST. (Feb. 4, 2020), https://knightcolumbia.org/content/from-private-bads-to-public-goods-adapting-public-utility-regulation-for-informational-infrastructure [https://perma.cc/5Y5Z-KA4Z]; *see also* Cohen, *supra* note 120.

be even counterproductive since these efforts tend to occlude the urgent need for meaningful regulation.