May 6, 2025

The Implications of Section 230 for Black Communities

The Impact of Section 230 Reforms on Black Communities

Danielle A. Davis, Esq.
Director, Tech Policy

JOINT CENTER FOR POLITICAL AND ECONOMIC STUDIES

Introduction

This series of issue briefs provides a summary of ***"The Implications of Section 230 for Black Communities"*** law review article that examines Section 230 of the Communications Decency Act (Section 230), which grants online platforms immunity from liability for user-generated content. The first brief explores the dual-edged nature of the statute, analyzing both the opportunities and challenges it presents for Black communities by protecting tech companies from liability for third-party content. The first brief highlights how this immunity from liability for third-party content empowers Black communities by enabling free expression and fostering innovation, while also examining how the same immunity allows harmful practices to persist. The second brief examines the nuances and legal implications of the statute, specifically how it affects platform moderation practices, including potential biases in content moderation that often disproportionately impact Black users. This third brief explores the implications of various Section 230 reform proposals with a focus on their unique impact on Black communities ■

The debate over Section 230 of the Communications Decency Act has become increasingly polarized, with Democrats typically advocating for stronger content moderation to combat disinformation and hate speech, while Republicans argue that platforms engage in political bias when moderating content, pushing for less restrictive rules.[1] Amid these debates, surveys suggest that Black Americans are slightly more inclined than other groups to support the ability to sue platforms for harmful actions over by third-party users.[2]

While Section 230 does not fully meet the needs of internet users today, fully repealing Section 230 is not a reasonable solution.[3] Doing so could stifle political activism, entrepreneurship, and free expression, particularly in Black communities, while simultaneously worsening the very issues Section 230 seeks to address, such as disinformation, discrimination, and hate speech. Yet failing to amend Section 230 enables the continuation of harmful content moderation practices that allow illegal harassment and the organization of white supremacist groups to thrive online.

Furthermore, potential reforms to Section 230 will not automatically hold platforms liable for all content. Plaintiffs would still need to meet legal standards, and courts could allocate damages between platforms and other responsible parties.[4] Reforms could also encourage over-moderation, as platforms may respond by excessively restricting content to avoid legal liability.[5] Therefore, any proposed changes must be carefully designed to avoid unintended consequences that could negatively impact Black communities.

While courts have ruled that platforms lose Section 230 protections when their algorithms or design features contribute to illegal discrimination,[6] some companies continue to act as though they are fully immune. With the rise of generative Artificial Intelligence (AI) and other advanced technologies, this challenge will only continue to grow.[7]

Therefore, reforms must make it clear that platforms are not protected when their algorithms or data practices contribute to unlawful discrimination. In the absence of legislative changes, courts should apply the "material contribution" test to determine how platform design and algorithms enable illegal activities that disproportionately harm Black communities.

Each potential reform has its pros and cons, and the best path forward may involve a combination of solutions. However, even reforms designed to benefit Black communities could face backlash, resulting in federal or state policies that weaken content moderation and amplify harmful content, including hate speech and white supremacy, targeting Black people.

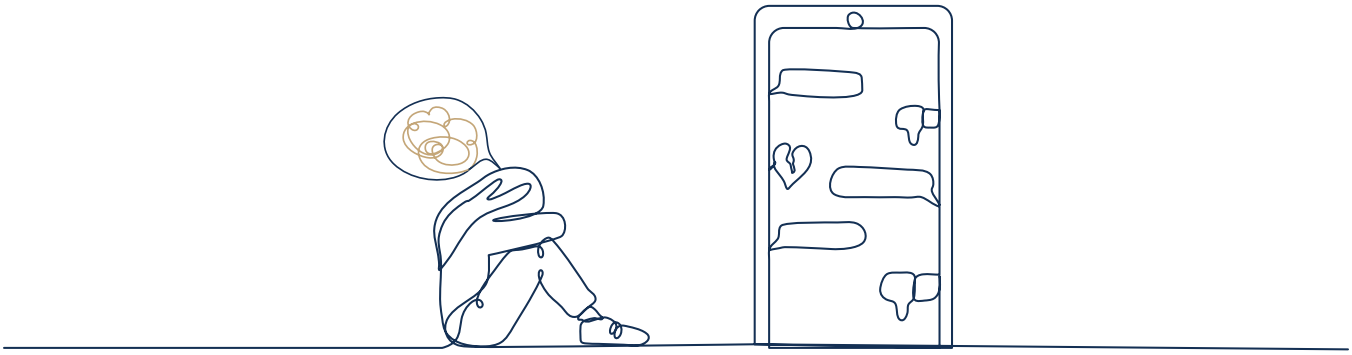
This issue brief will provide insights to help policymakers evaluate reform proposals based on how effectively they address the unique challenges facing Black communities, mitigate new risks, and avoid over-moderation that could hinder opportunities for Black creators and activists. In short, this issue brief will explore the implications of various Section 230 reform proposals with a focus on their unique impact on Black communities.

Civil Rights Carve-Outs

One potential solution is to create a carve-out in Section 230 to enforce civil rights laws against online platforms.[8] Section 230 already includes exceptions for federal criminal law, intellectual property law, and federal sex trafficking laws, so adding a civil rights carve-out would be consistent with existing practices.[9]

The fight against discrimination for Black communities is just as critical as for other groups benefiting from a civil rights carve out. Such a carve-out would hold social media platforms accountable for discriminatory practices, such as housing ads excluding Black users or biased treatment of Black guests on rental and ride-sharing platforms. It would reinforce the principle that platforms facilitating discrimination should not be shielded by Section 230 protections. Allowing platforms to profit from discriminatory ads—while being illegal in print under the Fair Housing Act—is indefensible.

However, defining the civil rights carve-out presents challenges.[10] Would it encompass laws such as the Fair Credit Reporting Act or state laws that prevent the spread of false election information, which is often used to suppress Black voters? The complexity of varying state laws could make it difficult to define violations, possibly prompting platforms to over-moderate content on race to avoid legal risks. Additionally, reliance on AI moderation could unintentionally suppress crucial content, such as activism exposing discrimination.



Some civil rights laws may also conflict with the interests of Black communities. For example, a platform could face lawsuits for hosting content about structural racism or The 1619 Project if a state such as Florida deems these topics to violate civil rights laws.[11]

One way to mitigate this risk is to limit the carve-out to specific federal statutes, such as the Voting Rights Act or the Ku Klux Klan Act.[12] These laws, along with provisions like those in the Violence Against Women Act[13] targeting hate crimes, could hold platforms accountable for facilitating violence or suppressing voting rights.

While this narrower carve-out might not address housing, employment, or financial discrimination, it would remove Section 230 immunity when platforms' designs, data, or algorithms contribute to illegal acts such as discriminatory ad targeting.

Yet this limited approach may leave other issues, such as racial harassment or illegal firearm sales, unaddressed. Additional carve-outs could be added,[14] but multiple exemptions might complicate compliance, increase litigation costs, and lead to the over-moderation of Black users' content.[15]

Algorithmic Recommendations Carve-Outs

Another potential carve-out to Section 230 could target platforms using algorithms to deliver and amplify content,[16] addressing the several challenges faced by Black communities. Platforms often use algorithms to drive engagement and profits, which can result in discriminatory practices, such as directing housing and job ads to white users while excluding Black users.[17]

Algorithms also have the potential to amplify disinformation, hate speech, and white supremacist content.[18] Since these algorithms are proprietary and hidden,[19] Black users cannot fully understand how their online experiences are shaped—an issue compounded by Section 230 protections.

While some algorithm-driven practices may already fall outside of Section 230 immunity when they materially contribute to illegal activities such as discriminatory ad targeting, explicitly exempting algorithmic decisions from Section 230 would ensure a consistent legal standard. This would hold platforms accountable, not for the content of third-party ads, but rather for using algorithms that promote discrimination, such as steering housing ads away from Black users.[20] Without this exemption, platforms are incentivized to profit from illegal discrimination and anti-Black practices.

Still, a broad carve-out for algorithmic content could also lead to unintended consequences. Platforms might reduce their investment in algorithms that benefit Black communities, citing concerns about increased litigation costs and the risks of third-party content. As a result, Black creators, businesses, and activists could lose crucial connections to their audiences, potentially replaced by lawyer-approved, safer content.[21] This carve-out could further hinder Black users' access to vital information by disrupting algorithms that currently prioritize relevant search results, making it harder for them to find tailored content and resources that address their specific needs and interests.

Additionally, stripping platforms of Section 230 immunity for algorithm-driven content could hinder the effectiveness of algorithms used to combat hate speech, white supremacy, discrimination, and other anti-Black activity.[22] To address these issues, the carve-out should be narrowly tailored to apply only to platforms whose algorithms materially contribute to illegal activities. [23]

Rather than blaming the technology itself, accountability should focus on those who use it unlawfully. While some argue that current legal interpretations already exclude algorithms from Section 230 when they contribute to illegal acts, a statutory clarification would ensure courts consistently apply the material contribution test.

Importantly, an algorithmic recommendation carve out would not address lawful but harmful content, such as hate speech, which algorithms may amplify. It also raises constitutional concerns, particularly in defining terms such as "algorithm" and "amplification."[24]

Ad Carve-Outs

Another reform proposal suggests removing Section 230 immunity for paid advertisements, commonly referred to as the "ad carve-out."[25] This proposal would clarify that platforms are not immune from liability for economic discrimination in housing, employment, and financial services ads that are directed toward white users and exclude Black users. Black consumers are often disproportionately targeted by deceptive practices, such as payday lending schemes and student debt relief fraud.[26] By implementing the ad carve-out, platforms would be encouraged to better vet their advertisers and prevent their use in fraudulent or discriminatory activities.

The reform would strip platforms of immunity, whether discrimination arises from biased advertisers using platform tools for racial targeting or algorithms and data collection that deliver ads discriminatorily without advertisers' knowledge. It also addresses ads that target Black voters with false information aimed at suppressing voter turnout, such as misleading boycott campaigns.[27]

In 2020, digital advertising accounted for 63 percent of total ad spending in the United States, and that figure is expected to rise to 75 percent by 2024.[28] Major players including Google, Meta, and Amazon dominate the digital ad market, holding more than 60 percent of ad revenue in 2022.[29] While traditional media, such as The New York Times, are legally restricted from publishing discriminatory housing ads, platforms such as Craigslist, Facebook, and Google are currently immune to liability for similar practices.[30] Holding online platforms to the same standard of care as traditional publishers for paid content would help address these disparities.

Section 230 protects platforms from liability for user-generated content due to the overwhelming volume of content,[31] but monitoring ads is more feasible given the smaller volume of ads. Therefore, it is reasonable to expect platforms to monitor paid ads and prevent profit from discriminatory practices.

Although removing Section 230 immunity for paid advertisements would address some challenges Black communities face, it would not solve problems such as anti-Black harassment, hate speech, or white supremacist organizing. It also wouldn't stop discriminatory practices on platforms including job boards, short-term rentals, or ride-sharing apps. Platforms could still profit from harmful anti-Black content and ads.

There are potential costs to this reform. If courts interpret the ad carve-out too broadly, it could expose Black-operated websites, such as small business sites or blogs that host third-party ads through services such as Google AdSense to liability. Stricter ad regulations also could also make it harder for Black communities to access vital information, such as health updates or voter mobilization efforts, due to increased costs and reduced ad distribution.

Notice and Takedown Proposals

Notice-and-takedown reform requires platforms to remove illegal content within a specified timeframe after being notified, or risk losing legal immunity. This process is already in place for copyright infringement in the United States.[32] and for illegal content in the EU, New Zealand, and South Africa,[33] as well as defamatory content in the UK.[34]

This section explores how a similar reform, specifically under the Internet Platform Accountability and Consumer Transparency Act (PACT Act), could affect Black communities. The PACT Act requires platforms to remove illegal content within four days of notice[35] or lose Section 230 immunity if they are aware of the content.[36] Illegal content includes material that federal or state courts have determined violates federal criminal or civil laws or state defamation laws. Platforms that fail to comply could lose their immunity for the flagged content.

Federal civil laws encompass crucial protections for Black Americans, such as civil rights and consumer protection laws. Platforms would no longer be able to use Section 230 immunity to avoid accountability for discriminatory housing ads that violate the Fair Housing Act, for example.[37] This reform could also target smaller, fringe platforms that serve as safe havens for white supremacist content, stripping them of immunity when they host illegal, anti-Black content in violation of federal laws.

Unlike other proposals, the notice-and-takedown approach would likely impose fewer costs on platforms, reducing the risk of service cutbacks or over-moderation that could negatively impact Black activists, entrepreneurs, or creators. Platforms would only be held responsible for content they are made aware of and fail to remove within the designated timeframe, instead of being liable for all unlawful content on their platform.

Yet this proposal does not fully address the challenges Section 230 presents for Black communities. Black users may lack the resources to secure a court order proving that content violates federal or state laws, which is required under this reform. Additionally, bad actors could exploit the system to silence Black voices. For instance, a platform might take down a post advocating "Save Black Lives – De-Unionize the Police" due to a baseless complaint from white supremacists claiming the content violates federal law. Platforms might find it easier to comply with such takedown requests than challenge them.

Moreover, this reform does not remove Section 230 immunity for "lawful-but-harmful" content, such as anti-Black hate speech or white supremacist organizing.[38] Nor does it address violations of state laws outside of defamation or federal law, such as deceptive political ads spreading disinformation about voting procedures.[39]

Content Neutrality Proposals

Content neutrality proposals are based on claims that tech companies selectively enforce guidelines to censor certain content, such as disinformation about election fraud claims or alternative COVID-19 treatments, hate speech such as dehumanizing language about Black, and immigrant communities, and incitement to violence such as "stop the steal" posts tied to the January 6, 2021 attack on the U.S. Capitol. Proponents often argue that these reforms are necessary to protect conservative speech from suppression by tech companies. [40]

Some proposals in Congress aim to remove the phrase "otherwise objectionable" from the types of content platforms can moderate under Section 230(c)(2).[41] These reforms seek to limit platforms' ability to moderate content, which poses significant challenges because much harmful content—such as hate speech, voter suppression, and disinformation—may not fit into categories such as "obscene" or "violent."[42] As a result, Black communities could face increased exposure to unchecked harmful content. Additionally, platforms including short-term rental and ride-sharing apps could lose the ability to remove users engaging in racial discrimination.

Other content-neutrality reforms propose removing Section 230 immunity for platforms that restrict speech, [43] including proposals to ban moderation of all content except illegal material[44] or target platforms for perceived political bias.[45] States including Texas and Florida have passed laws limiting platforms' moderation of user viewpoints[46] or prohibiting the de-platforming of political candidates.[47]

However, the argument that content moderation stifles free speech misunderstands the First Amendment, which applies to the government, not private companies.[48] Social media platforms, as private entities, have the right to moderate content under their terms, and users agree to follow these community standards, which often prohibit hate speech, extremist content, and misinformation; or engaging in impersonation, bullying, or harassment.

Forcing platforms to treat all lawful content equally would exacerbate problems for Black communities, allowing hate speech, white supremacist content, and disinformation to spread. Currently, platforms have the discretion to remove harmful content, but content neutrality reforms would restrict their ability to do so.

Section 230 was originally designed to allow platforms to remove harmful content without facing legal liability.[49] Limiting this ability undermines protections that benefit Black communities by keeping dangerous content off of online platforms.[50]

While some suggest that Black communities should support content-neutrality reforms due to disproportionate de-platforming, these proposals do not address the specific issues Black users face, such as the unjust removal of their content, which often occurs even when it complies with platform guidelines.[51]

Size-based Carve-Outs and Disclosure Requirements

Several proposed reforms to Section 230 can be combined with additional reforms,[52] such as size-based exemptions or disclosure mandates. A size-based exemption could apply to platforms with more than five million users or $100 million in annual revenue, focusing on larger companies that play a significant role in spreading harmful content and have the resources to comply with new regulations. These exemptions are often included in reform proposals related to civil rights, algorithms, advertising, notice-and-takedown proposals, and content-neutrality proposals.[53]

Size-based exemptions would allow small, Black-owned startups to innovate without facing heavy regulatory burdens while ensuring more oversight for larger social media and sharing-economy platforms where Black users are highly active.[54] But this approach has a downside: many threats to Black communities originate from smaller platforms such as 8chan, Gab, and Parler, all of which serve as hubs for white supremacy. Exempting these platforms could allow Section 230 protections to continue shielding harmful activities on these sites.[55]

Disclosure reforms could require platforms to regularly share information about their content moderation standards and efforts, often as part of other Section 230 reforms.[56] Such mandates would encourage platforms to engage in responsible moderation of harmful content such as hate speech, white supremacy organizing, and disinformation, without requiring government-imposed removal, which could raise constitutional concerns.[57] Transparency would also shed light on algorithms and content practices, helping Black communities and policymakers better understand how discrimination, hate speech, and disinformation spread under Section 230 protections.[58]

Excessive disclosure requirements, however, could deter platforms from moderating harmful content targeting Black communities. If platforms were required to provide detailed justifications for every removed or downranked post, along with potential penalties for content removal, then this could stifle effective moderation. Still, large platforms, particularly those already complying with regulations under the EU's Digital Services Act, could adapt to these requirements without significant burdens, reducing the risk of over-moderation.[59]

Conclusion

Section 230 reform offers both opportunities and risks for Black communities. While reforms such as civil rights, algorithmic, and ad carve-outs could hold platforms accountable for discriminatory practices, they must be carefully designed to avoid stifling political activism, entrepreneurship, and free expression, especially for Black creators and small businesses.

Size-based exemptions and disclosure mandates can ensure larger platforms with significant influence are held accountable, while allowing smaller, Black-owned startups to grow without excessive regulation. But fringe platforms that promote hate speech, including Gabe, 8chan, and Parler, should not be exempt from these reforms.

Reforms such as notice-and-takedown and content-neutrality proposals must be evaluated for their real-world impacts on disinformation, hate speech, and platform accountability. A balanced approach is crucial to ensure reforms protect Black communities from harm while supporting their ability to use digital platforms for advocacy and justice.

Endnotes

1. *See* Ashley Johnson and Daniel Castro, *Proposals to Reform Section 230*, ITIF (Feb. 22, 2021) (explaining traditional positions of President Trump and most conservatives, and President Biden and those on the left).

2. *See* Emily A. Vogels, *Americans' Views on How Online Harassment Should Be Addressed*, Pew Research Ctr. (Jan. 13, 2021) (finding that 45% of Black adults favor the ability of online harassment victims to sue platforms where harassment occurred, compared with 41% of Latinx and Asian adults and only 28% of white adults); Data for Progress, *National Poll*, Vox, at 4 (Jan. 2021) (indicating that 44% of Black Americans, 45% of Latinos, and 38% of whites think that tech companies should be more legally liable for content users post on their sites). *But see National Tracking Poll #2112020*, Morning Consult, Dec. 3–7, 2021, at 156–157 (finding that a larger share of Black Americans (37%) than white Americans (31%) *strongly support* allowing social media companies to be held at least somewhat liable in courts and lawsuits for the actions of their users, but that a larger share of white Americans (68%) than Black Americans (58%) *either strongly support or somewhat support* allowing social media companies to be held at least somewhat liable in courts and lawsuits for the actions of their users).

3. *See* Abandoning Online Censorship Act, H.R. 8896, 116th Cong. (2020) (legislation introduced that would repeal Section 230); A Bill to Repeal Section 230 of the Communications Act of 1934, S. 2972. 117th Cong. (2021) (legislation introduced that would repeal Section 230).

4. *See* Brief of Amici Curiae of The Cyber Civil Rights Initiative and Legal Scholars in Support of Petitioners, *Gonzalez v. Google LLC*, 143 S. Ct. 762 (Dec. 7, 2022), No. 21-1333, at 5 [hereinafter Cyber Civil Rights Initiative Amicus Brief (*Gonzalez*)] ("[T]he absence of immunity is not synonymous with the presence of liability.").

5. *See* Johnson and Castro, *Proposals to Reform Section 230*, *supra* note 1, ITIF (Feb. 22, 2021) ("Instead of just removing content that clearly violates the law or their terms of service, they would also likely remove any content that falls into a gray area… because to not do so would mean risking legal trouble.").

6. *See Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1167–68 (9th Cir. 2008) ("a website helps to develop unlawful content, and thus falls within the exception to section 230, if it *contributes materially* to the alleged illegality of the conduct"); *Vargas v. Facebook, Inc.*, No. 21-16499, 2023 WL 4145434, at *2- 3 (9th Cir. June 23, 2023) (rejecting Facebook claims that Section 230 immunized it from claims for discriminatory ad distribution); *Lemmon v. Snap*, 995 F.3d 1085, 1093 (2021) (refusing to apply Section 230 immunity to a negligent design lawsuit for an application that encouraged users to drive at high speeds and post the speed).

7. *See, e.g.*, No Section 230 Immunity for AI Act, S. 1993, 118th Cong. (2023) (waiving immunity under Section 230 for claims and charges related to artificial intelligence).

8. *See, e.g.*, The Civil Rights Modernization Act of 2021, H.R. 3184, 117th Cong. (2021) (removing Section 230 immunity for platforms for ads targeted with algorithms that violate civil rights laws); Safeguarding Against Fraud, Exploitation, Threats, Extremism, and Consumer Harms (SAFE TECH) Act, S. 560, 118th Cong. (2023) (removing 230 protections as a defense in various cases, including cases related to ads and cases alleging violations of civil rights, harassment, or intimidation laws (including those based on race)); Olivier Sylvain, *Discriminatory Designs on User Data; Exploring How Section 230's Immunity Protections May Enable or Elicit Discriminatory Behaviors Online*, Knight First Amend. Inst. Colum. U. (Apr. 1, 2018) ("There is no reason why Congress couldn't also write in an explicit exception to Section 230 immunity for violations of civil rights laws."); Spencer Overton, *State Power to Regulate Social Media Companies to Prevent Voter Suppression*, 53 U.C. Davis L. Rev. 1793, 1827 (2020) ("Congress should explicitly acknowledge that Section 230 does not provide a defense to federal and state civil rights claims arising from online ad targeting.").

9. 47 U.S.C. § 230(e)(1)–(5).

10. The problem of lack of clarity in scope also exists under the current exceptions for intellectual property law (which would include state law related to trademarks and trade secrets) and state laws similar to the federal Electronic Communications Privacy Act of 1986. 47 U.S.C. § 230(e)(2)&(4). *See* Johnson and Castro, *Proposals to Reform Section 230*, *supra* note 1, ITIF (Feb. 22, 2021).

11. Individual Freedom Act, Florida H.B. 7 (2022); *see also* Janai Nelson, Op-Ed: *Ron DeSantis Wants to Erase Black History. Why? The New York Times* (Jan. 31, 2023).

12. Some proposed Section 230 reforms have limited the scope of the civil rights exemption to a couple of federal civil rights provisions. *See, e.g.*, Protecting Americans From Dangerous Algorithms Act, H.R. 2154, 117th Cong. (2021) (removing Section 230 immunity in civil cases involving particular federal civil rights violations (Sections 1980 and 1981) or particular federal terrorism violations if platforms use algorithms to disseminate and amplify content at issue).

13. VAWA (1994); 2022 VAWA Reauthorization.

14. *See e.g.*, Accountability for Online Firearms Marketplaces Act of 2021, S. 2725, 117th Cong. (2021) (removing Section 230 immunity for online firearms marketplaces, including those that facilitate firearms transactions or ads, or make available digital instructions to program a 3D printer to produce a firearm).

15. *See* Carol M. Rose, *Crystals and Mud in Property Law*, 40 Stan. L. Rev. 577, 578–79 (1988) (discussing phenomenon when "crystalline rules have been muddied repeatedly by exceptions" so that parties don't know their rights).

16. *See, e.g.*, Platform Integrity Act, H.R. 9695, 118th Cong. (2022); Justice Against Malicious Algorithms Act of 2021, H.R. 5596, 117th Cong. (2021); Don't Push My Buttons Act, H.R. 8515, 116th Cong. (2020); Protecting Americans From Dangerous Algorithms Act, H.R. 2154, 117th Cong. (2021). The lack of Section 230 protections for third-party content amplified by algorithms does not equate to automatic liability—a plaintiff would still need to establish the elements of the underlying legal claim against the platform. *See, e.g., Twitter, Inc. v. Taamneh*, 598 v. 471 (2023) (finding that plaintiffs did not show that Twitter gave such knowing and substantial assistance to ISIS that they consciously and culpably participated in the alleged tort stemming from the terrorist attack, and that recommendation algorithms did not convert Twitter's passive assistance into active abetting).

17. *See, e.g.*, Safiya Umoja Noble, Algorithms of Oppression, 173–79 (2018) (detailing the story of a Black hairdresser whose business was diminished by Yelp's practices of discarding reviews of her customers who were infrequent Yelp users and charging her fees to remove the links to her competitors on her business's Yelp page); Safiya Noble, *Google Has a Striking History of Bias Against Black Girls*, *Time* (Mar. 26, 2018).

18. *See e.g., Id.*

19. Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (2016); Susan Benesch, *Nobody Can See Into Facebook*, *The Atlantic* (Oct. 30, 2021).

[20] Daphne Keller, *Amplification and Its Discontents*, Knight First Amendment Inst. at Columb. Univ., June 8, 2021 ("I have argued elsewhere that discrimination claims like this are likely not properly subject to intermediary liability immunities under CDA 230. . . . The problem in those cases is . . . that platforms introduce harm distinct from that content through their ranking or targeting."); Cyber Civil Rights Initiative Amicus Brief (*Gonzalez*), *supra* note 4 (emphasizing the importance of not expanding 230 immunity to discriminatory decisions such as "recruiting algorithms that discriminatorily and unlawfully screen women from job opportunities, mortgage approval algorithms that disproportionately and unlawfully reject applications on the basis of race, and facial recognition systems that produce inaccurate matches on the basis of race or sex"). In the oral argument *in Gonzalez v. Google*, Google's attorney conceded to Supreme Court Justice Barrett that algorithmic race-based discrimination is not immunized by Section 230 because the discrimination turns on the website's conduct rather than the third-party speech. Transcript of Oral Argument, at 140–142, Gonzalez v Google, LLC, 143 S. Ct. 762 (2023) (No. 21-1333).

[21] Civil Rights Scholars Amicus Brief (*Gonzalez*), *supra* note 4 ("If platforms are liable for the third-party content they recommend, they will only be willing to recommend 'safe' content.... [C]ontent creators from underserved and marginalized groups would be crowded out.").

[22] *See* Cyber Civil Rights Initiative Amicus Brief (*Gonzalez*), *supra* note 4, at 6–7 (arguing that categorical denial of immunity to platforms using targeted algorithms could dissuade such platforms from using algorithms to remove or reduce the accessibility of harmful material); Civil Rights Scholars Amicus Brief (*Gonzalez*), *supra* note 4 ("If platforms may be liable for recommending content using a neutral algorithm, it follows that they could be liable for moderating content by the same means."); *The Role of AI in Addressing Misinformation on Social Media Platforms*, Ctr. for Data Ethics and Innovation, at 18–20 (2021).

[23] *See, e.g.*, Justice Against Malicious Algorithms Act of 2021, H.R. 5596, 117th Cong. (2021) (removing Section 230 immunity when platforms knowingly or recklessly use an algorithm to make a personalized recommendation of third-party information that "materially contributed to a physical or severe emotional injury to any person").

[24] *See, e.g.*, Bertram Lee, *Where the Rubber Meets the Road: Section 230 and Civil Rights*, Public Knowledge (Aug. 12, 2020)

[25] Safeguarding Against Fraud, Exploitation, Threats, Extremism, and Consumer Harms (SAFE TECH) Act, S. 560, 118th Cong. (2023). Other proposals would remove immunity for platforms that sell targeted ads, even if the claim does not arise from speech within a particular targeted ad. *See* Break Up Big Tech Act of 2020, H.R. 8922, 116th Cong. (2020); Behavioral Advertising Decisions Are Downgrading Services (BAD ADS) Act, S. 4337, 116th Cong. (2020). Another proposal would ban political microtargeting completely. Banning Microtargeted Political Ads Act, H.R. 7014, 116th Cong. (2020).

[26] *See* Federal Trade Commission, Serving Communities of Color: A Staff Report on the Federal Trade Commission's Efforts to Address Fraud and Consumer Issues Affecting Communities of Color, at 3, 40, 43 (Oct. 2021).

[27] *See, e.g.*, Russian Trolls' Chief Target Was 'Black US Voters' in 2016, BBC (Oct. 9, 2019).

[28] Vision Monday Staff, *Forecast Shows Digital Ad Spending Jumping 25 Percent in 2021 as Economy Recovers,* Vision Monday (Apr. 16, 2021) (citing eMarketer data and projections).

[29] Statista Research Department, *Share of Ad-Selling Companies in Digital Advertising Revenue in the United States from 2020 to 2025*, Statista (May 30, 2023).

[30] *See Civil Rights and Discrimination in Housing, Employment, and Credit* discussion in issue brief two.

[31] *See* Pauline T. Kim, *Manipulating Opportunity*, 106 Va. L. Rev. 867, 926–27 (2020) (examining Section 230(c)(1)'s facilitation of racial and gender inequality in labor markets).

[32] The Digital Millennium Copyright Act (DMCA), 17 U.S.C. § 512 (c)–(d) (1998).

[33] The Digital Services Act, Commission Regulation 2022/2065 Art. 22 (E.U.) (extending immunity to providers who, "upon obtaining actual knowledge of ...illegal activities or content, act expeditiously to remove or to disable access to that content..."); Harmful Digital Communications Act 2015, s. 6 (N.Z.); Electronic Communications and Transactions Act of 2002 Chapter XI, Section 77 (S. Afr.); *See* Ashley Johnson & Daniel Castro, *How Other Countries Have Dealt With Intermediary Liability*, ITIF (Feb. 22, 2021).

[34] Defamation Act 2013 § 5 (U.K.) (creating notice and takedown procedures for defamation).

[35] Internet Platform Accountability and Consumer Transparency (Internet PACT) Act, S. 483, 118th Cong. § 5 (2023).

[36] *Id.*

[37] *See, e.g., Chicago Lawyers' Committee for Civil Rights Under Law, Inc. v. Craigslist, Inc.*, 519 F.3d 666, 668 (7th Cir. 2008).

[38] The PACT Act does establish a process for filing complaints about content that does not comply with platform guidelines and directs platforms to review the complaints and take "appropriate steps." Internet Platform Accountability and Consumer Transparency (Internet PACT) Act, S. 483, 118th Cong. § (5)(c)(1)(B) (2023).

[39] Including more state law violations, however, could also lead to unintended consequences, such as Florida suing a platform for posting an employer's diversity training on systemic racism that violates the technical terms of a new state law. 2022 Fla. Laws Ch 2022-72, *amending* Fla. Stat. § 1003.42.

[40] *See, e.g., Preserving Free Speech and Reining in Big Tech Censorship: Hearing Before the H. Comm. on Energy and Commerce, Subcomm. on Comms. & Tech.*, 118th Cong. (Mar. 28, 2023).

[41] *See, e.g.,* Disincentivizing Internet Service Censorship of Online Users and Restrictions on Speech and Expression (DISCOURSE) Act, S. 921, 118th Cong. (2023) (amending 230(c)(2) by removing "otherwise objectionable" and removing 230 protections for "dominant market share" platforms that reasonably appear to promote a discernable viewpoint or suppress legitimate speech); Stop the Censorship Act, H.R. 8612, 118th Cong. (2022) (amending 230(c)(2) by allowing platforms to use the provision as a defense only when they moderate unlawful material, and removing immunity when platforms moderate offensive but legal material); 21st Century Free Speech Act, H.R. 7613, 118th Cong. (2022) (completely repealing Section 230 and replacing it with a Section 232, and removing immunity from any platform that increases or decreases the dissemination or visibility of third-party material, except material that is obscene, lewd, lascivious, filthy, excessively violent, harassing, promoting self-harm, or unlawful); Preserving Political Speech Online Act, S. 2338, 117th Cong. (2021) (amending 230(c)(2) by limiting the reasons for "good faith" removal of material to content that is obscene, illegal, excessively violent, harassing, threatening, or promoting illegal activity).

[42] *See* 47 U.S.C. § 230(c)(2).

[43] *See* Danielle Keats Citron & Mary Anne Franks, *The Internet As a Speech Machine and Other Myths Confounding Section 230 Reform*, 2020 U. Chi. Legal F. 45, 61–67 (2020) (detailing the arguments that equate platform content moderation with impermissible censorship under the First Amendment, and explaining the shortcomings of the arguments).

[44] *See, e.g.,* Stop the Censorship Act, H.R. 8612, 118th Cong. (2022).

[45] *See, e.g.*, Curtailing Online Limitations That Lead Unconstitutionally to Democracy's Erosion (COLLUDE) Act, S. 1525, 118th Cong. (2023); Disincentivizing Internet Service Censorship of Online Users and Restrictions on Speech and Expression (DISCOURSE) Act, S. 921, 118th Cong. (2023); Ending Support for Internet Censorship Act, S. 1914, 116th Cong. (2019).

[46] Texas House Bill 20, 87th Session (2021); *see NetChoice, L.L.C. v. Paxton*, 49 F.4th 439, 455, 494 (5th Cir. 2022 (upholding Texas statute mandating viewpoint neutral content moderation by social media platforms).

[47] Florida Senate Bill 7072 (2021); *see NetChoice, LLC, v. Att'y Gen., Fla.*, 34 F.4th 1196, 1231 (11th Cir. 2022) (granting a preliminary injunction because it was substantially likely that a Florida statute's regulation of social media content moderation was unconstitutional).

[48] *Manhattan Cmty. Access Corp. v. Halleck,* 139 S. Ct. 1921, 1926 (2019) ("[t]he Free Speech Clause of the First Amendment constrains governmental actors and protects private actors").

[49] Cyber Civil Rights Initiative Amicus Brief (*Gonzalez*), *supra* note 4, at 8–11 (reviewing the legislative history of Section 230 and establishing that "the crafters of Section 230 did not seek to relieve ICSPs of any responsibility for harmful content appearing on their platforms. On the contrary, they intended Section 230 to enable and incentivize ICSPs to moderate content to protect users from harm.").

[50] *See* Brian Fishman, *Dual-Use Regulation: Managing Hate and Terrorism Online Before and After Section 230 Reform*, The Brookings Institution (Mar. 14, 2023).

[51] *See supra* text accompanying notes 225–229 [insert last]

[52] Various other reform proposals exist. Scholars Danielle Citron and Benjamin Wittes, for example, propose limiting Section 230 immunity to a provider that "takes reasonable steps to prevent or address unlawful uses of its services." Danielle Keats Citron & Benjamin Wittes, *The Internet Will Not Break: Denying Bad Samaritans Section 230 Immunity*, 86 Fordham L. Rev. 401, 419 (2017). This proposal would encourage companies to take reasonable steps to design their platforms, data collection procedures, and algorithms in ways that avoid unlawful discrimination in housing, employment, lending, voting, and other contexts. The proposal would not, however, stop the Section 230 immunity subsidy to "lawful-but-awful" activities that harm Black communities, such as hate speech, non-violent white supremacy organizing, digital blackface, medical disinformation, and domestic political operatives' deceptive ads targeted to discourage Black users from voting.

[53] *See, e.g.*, Protecting Americans From Dangerous Algorithms Act, H.R. 2154, 117th Cong. (2021); Justice Against Malicious Algorithms Act of 2021, H.R. 5596, 117th Cong. (2021); Internet Platform Accountability and Consumer Transparency (Internet PACT) Act, S. 483, 118th Cong. § 5(e) (2023); Behavioral Advertising Decisions Are Downgrading Services (BAD ADS) Act, S. 4337, 116th Cong. (2020); 21st Century Free Speech Act, H.R. 7613, 118th Cong. (2022); Florida Senate Bill 7072 (2021); Texas House Bill 20 (2021). *See also Digital Services Act: Commission Designates First Set of Very Large Online Platforms and Search Engines*, European Commission (Apr. 25, 2023) (designating 17 platforms as Very Large Online Platforms ("VLOPs")—*e.g.*, Amazon, Facebook, Instagram, LinkedIn, X, YouTube—and 2 platforms as Very Large Online Search Engines ("VLOSEs")—Google and Bing—and requiring that these platforms comply with new obligations under the Digital Services Act).

[54] Size-based carve-outs may create a disincentive for Black start-ups to grow beyond a certain size that would subject them to potential liability. One alternative is to apply reforms not to platforms of a minimum size (in terms of monthly users or revenue), but to content that reaches an audience of a certain size. Since content that goes viral is limited, moderating content based on attention metrics would be more manageable, would focus on the most impactful content, and would not significantly deter platform growth. Special thanks to Harvard lecturer and former YouTube engineer Hong Qu for flagging attention metrics.

[55] *See* Johnson & Castro, *Proposals to Reform Section 230*, *supra* note 1, ITIF (Feb. 22, 2021). ("[T]here are smaller online services that profit directly from illegal or abusive third-party content. . . and under a size-based carve-out, they would continue to benefit from Section 230 immunity. . . .").

[56] *See, e.g.*, Internet Platform Accountability and Consumer Transparency (Internet PACT) Act, S. 483, 118th Cong. § 5(e) (2023); Florida Senate Bill 7072 (2021); Texas House Bill 20 (2021). Some proposed regulations focus on disclosure. *See, e.g.*, Platform Accountability and Transparency Act, S. 5339, 118th Cong. (2022); Algorithmic Accountability Act of 2022, H.R. 6580, 117th Cong. (2022); Digital Services Act, Commission Regulation 2022/2065 Art. 15-17 (E.U.).

[57] Disclosure requirements are generally subject to less-exacting scrutiny by courts than affirmative limitations on speech. *See, e.g.*, Milavetz, Gallop and *Milavetz, P.A. v U.S.*, 559 U.S. 229, 250 (2010); Citizens United v. Federal Election Comm'n, 558 U.S. 310, 366–67, 369 (2010).

[58] Pasquale, *supra* note 31; Benesch, *supra* note 31; Mark MacCarthy, *Transparency is Essential for Effective Social Media Regulation*, The Brookings Institution (Nov. 1, 2022).

[59] *NetChoice v. Florida*, 34 F.4th at 1230.

[60] Digital Services Act, Commission Regulation 2022/2065 Art. 15-17 (E.U.) (detailing transparency reporting obligations of platforms and notice and action mechanisms).

Acknowledgments

The Joint Center is especially appreciative of the Knight Foundation for supporting this report. We also thank Spencer Overton and Catherine Powell, who co-wrote the law review article this report summarizes.

A special thanks to Dedrick Asante-Muhammad, Jessica Fulton, Victoria Johnson, Kennedy Fortner, and Ed Paisley for their support throughout the project and for closely reviewing the report to ensure the highest level of accuracy. This report was designed by Vlad Archin.

Media Contact
press@jointcenter.org | 202.789.3500 EXT 105

Joint Center for Political and Economic Studies
633 Pennsylvania Ave., NW
Washington, DC 20004
info@jointcenter.org | jointcenter.org

@JointCenter
linkedin.com/company/jointcenter
facebook.com/JointCenter1970
@JointCenter
youtube.com/@jointcenterYT

JOINT CENTER
FOR POLITICAL AND ECONOMIC STUDIES