24-271

IN THE

United States Court of Appeals

FOR THE NINTH CIRCUIT

X CORP.,

Plaintiff-Appellant,

—v.—

ROBERT BONTA, IN HIS OFFICIAL CAPACITY AS ATTORNEY GENERAL OF CALIFORNIA,

Defendant-Appellee.

ON APPEAL FROM THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF CALIFORNIA

BRIEF FOR PLAINTIFF-APPELLANT

WILLIAM R. WARNE
MEGHAN M. BAKER
DOWNEY BRAND LLP
621 Capitol Mall, 18th Floor
Sacramento, California 95814
(916) 444-1000

JOEL KURTZBERG
FLOYD ABRAMS
JASON D. ROZBRUCH
LISA J. COLE
CAHILL GORDON & REINDEL LLP
32 Old Slip
New York, New York 10005
(212) 701-3120

Attorneys for Plaintiff-Appellant

DISCLOSURE STATEMENT

Pursuant to Rule 26.1 of the Federal Rules of Appellate Procedure, X Corp. states that it is a privately held corporation. Its parent corporation is X Holdings Corp. No publicly traded corporation owns 10% or more of the stock of X Corp. or X Holdings Corp.

DATED: February 14, 2024

/s/ Joel Kurtzberg

CAHILL GORDON & REINDEL LLP
Joel Kurtzberg
Floyd Abrams
Jason Rozbruch
Lisa J. Cole
32 Old Slip
New York, NY 10005
Telephone: 212-701-3120
Facsimile: 212-269-5420

DOWNEY BRAND LLP
William R. Warne
Meghan M. Baker
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Telephone: 916-444-1000
Facsimile: 916-444-2100

Attorneys for Appellant X Corp.

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Other Authorities

INTRODUCTION

California enacted Assembly Bill 587 ("AB 587")[1]—a state law compelling social media companies to provide the State with semi-annual disclosures about their efforts to moderate certain categories of constitutionally protected speech that the State disfavors—as part of a concerted effort to limit or eliminate those categories of speech on social media platforms. The California Legislature was clear about both its intent and approach: it intentionally picked the most controversial and difficult-to-define categories of "awful but lawful" content—hate speech, racism, extremism, radicalization, disinformation, misinformation, harassment, and foreign political interference. And it imposed an approach requiring the companies to define the categories (or decline to do so) and provide statistics to the Attorney General ("AG") about moderation of such content on their platforms as part of an effort to "pressure" the social media companies into restricting speech that the government finds objectionable or undesirable.

The law also provides nearly unfettered discretion to the AG to impose or threaten to impose substantial costs on social media companies—through costly document and other information requests and threatened or actual enforcement actions—if those companies fail to moderate these categories of content in the

---

[1] Codified at California Business & Professions Code §§22675-81.

manner that the State desires. This concern is not purely theoretical. AG Bonta has already sent a "threat letter" to X Corp. (formerly Twitter, Inc.) and other major social media companies stating that he will not hesitate to enforce AB 587 if the companies do not do more to address his concerns about "disinformation" and "misinformation" on their platforms. Both the structure of the law and the manner in which AG Bonta has already threatened to apply it impermissibly interfere with social media companies' rights under the First Amendment and the California Constitution to moderate user content on their platforms free of state interference with their editorial judgment.

The State attempts to justify AB 587 as a purported "transparency measure" that "merely" requires social media companies to provide the State with information about their content-moderation policies and how they are applied. But what the State euphemistically refers to as a "transparency measure" in fact has the purpose and effect of compelling X Corp. to take positions on some of the most controversial topics of the day—what constitutes hate speech, racism, misinformation, and so forth—and grants the AG the power to pressure X Corp. to modify its moderation decisions, on pain of monetary and injunctive sanctions, based on whether the AG deems X Corp.'s disclosures about its moderation activities to "materially omit[] or misrepresent[]" its practices. But this standard is so vague that it creates no standard at all. Instead, it impermissibly provides the AG unfettered discretion to impose

substantial costs on social media companies for content-moderation practices that simply are not to the AG's liking.

AB 587 thus violates X Corp.'s free speech rights because it compels X Corp. to engage in controversial speech against its will; impermissibly interferes with X Corp.'s constitutionally protected editorial judgments about what content may appear on its social media platform; has both the purpose and likely effect of pressuring X Corp. to disfavor, remove, or deprioritize constitutionally protected speech that the State deems undesirable; and places an unjustified and undue burden on X Corp. Moreover, AB 587 was intended to and does interfere with the immunity afforded to X Corp. under the Communications Decency Act ("CDA") Section 230 ("Section 230"), which protects X Corp.'s ability to self-regulate these categories of content in the manner that X Corp., not the State, sees fit.

The parties and the district court all acknowledged that AB 587 is a content-based statute because it "target[s] speech based on its communicative content"—a designation that, absent a historical exception to the bedrock general rule, renders laws "presumptively unconstitutional" and triggers strict scrutiny. *Reed* v. *Town of Gilbert, Ariz.,* 576 U.S. 155, 163 (2015). Yet, in denying X Corp.'s preliminary injunction motion, the district court applied the less-rigorous standard of review in *Zauderer* v. *Office of Disciplinary Couns.*, 471 U.S. 626 (1985), which applies only to compelled commercial disclosures that involve purely factual, uncontroversial

information about the terms under which services will be available. AB 587 defies such a description.

To begin, AB 587 compels non-commercial speech. The district court appeared to agree, observing throughout oral argument that it "[did not] think this is commercial speech here," 2-ER-174, a conclusion that in and of itself precludes application of *Zauderer*. The district court echoed this conclusion in its decision denying the preliminary injunction motion, holding that "[t]he reports to the Attorney General compelled by AB 587 do not so easily fit the traditional definition of commercial speech" because "[t]he compelled disclosures are not advertisements, and the social media companies have no particular economic motivation to provide them." 1-ER-5.

Although these findings should preclude application of *Zauderer*, the district court applied *Zauderer* anyway, primarily because the Fifth and Eleventh Circuits applied *Zauderer* to Texas and Florida statutes that, unlike AB 587, do not target any particular categories of content identified by the State, let alone controversial categories like hate speech, misinformation, extremism, and foreign political interference. 1-ER-5 (citing *NetChoice, LLC* v. *Att'y Gen., Fla.* ("*NetChoice (Fla.)*"), 34 F.4th 1196, 1230 (11th Cir. 2022), and *NetChoice, LLC* v. *Paxton* ("*NetChoice (Tex.)*"), 49 F.4th 439, 485 (5th Cir. 2022)).

In sum, the district court erred in applying *Zauderer* for each of the following

independent reasons: AB 587 (1) compels non-commercial speech; (2) compels disclosures that are not "purely factual, uncontroversial" information about the terms under which X Corp.'s services will be available; (3) does not further a substantial governmental interest; and (4) targets specific content and viewpoints in a manner the statutes addressed in the *NetChoice* cases do not.

Strict scrutiny must instead apply here because AB 587 (1) regulates speech based on its content and viewpoint; (2) allows the State to interfere impermissibly with X Corp.'s constitutionally protected editorial judgment about content on its social media platform; and (3) regulates "speech about speech" in a way that impacts multiple sets of speech rights—those of X Corp., those of users that post on X, and those of users that read posts on X.

In the end, AB 587 fails under any level of scrutiny. It does not directly and materially advance a substantial, important, or compelling government interest and limits more speech than is necessary to accomplish any legitimate governmental goal. The district court thus erred in holding that X Corp. was unlikely to succeed on its First Amendment and Article I, Section 2, claims.

The district court also erred in concluding that X Corp. was unlikely to succeed on its Section 230 claims. The district court rejected X Corp.'s Section 230 claims, holding that AB 587 "does not provide for any potential liability stemming from a company's content-moderation activities *per se*," even if it permits the AG

to threaten liability—as he has already done—as a way to pressure X to change its content-moderation practices. 1-ER-8. Because AB 587 has both the purpose and effect of allowing the State to pressure X Corp. to change its content-moderation policies and imposes liability for good faith content-moderation decisions unless they are made in the manner dictated by the State, AB 587 directly conflicts with Section 230, which protects X Corp.'s ability to self-regulate the content that appears on X.

JURISDICTIONAL STATEMENT

The district court had jurisdiction under 28 U.S.C. §§1331, 1367, and 1343(a), and 42 U.S.C. §1983. It denied X Corp.'s preliminary injunction motion on December 28, 2023.

X Corp. filed a timely notice of appeal on January 12, 2024. 6-ER-1094–95; Fed. R. App. P. 4(1)(A). This Court has jurisdiction under 28 U.S.C. §1292(a)(1).

STATUTORY AUTHORITIES

AB 587 and Section 230 are attached hereto as addenda.

ISSUES PRESENTED

I. Did the district court err by refusing to apply strict scrutiny, and instead applying *Zauderer*—a standard that applies only to compelled commercial disclosures consisting of purely factual, uncontroversial information about the terms under which services will be available—to AB 587, a law with the stated purpose of

pressuring social media companies to change their content-moderation policies to limit or remove content that the State disfavors?

II. Did the district court err by holding that AB 587's TOS Report survives First Amendment scrutiny, regardless of which level of scrutiny applies?

III. Did the district court err by holding that AB 587—which is designed to and does allow the State to pressure X Corp. to change its content-moderation policies if they are not to the State's liking—is not preempted by 47 U.S.C. §230(c)(2), which precludes the State from holding interactive computer service providers liable for good faith efforts to moderate objectionable content?

STATEMENT OF THE CASE

Nature of the Case

AB 587 Applies to Appellant

X Corp. operates X (formerly Twitter), which is a "social media platform" under §22675(e) because it (1) is a public internet-based application with users in California; (2) connects users to allow them to interact socially; (3) allows users to construct public or semi-public profiles to sign into and use the application; and (4) allows users to share a social connection with and post content viewable to others. 7-ER-1104–05. AB 587 applies to X Corp. because it generated more than $100 million in gross revenue in 2023. 7-ER-1104.

AB 587's Statutory Scheme

AB 587 has three main components: (1) a requirement that social media companies publicly post their terms of service, including processes for flagging content and potential actions that may be taken with respect to flagged content ("Terms of Service (or 'TOS') Requirement"), *see* §22676; (2) a requirement that social media companies submit to AG Bonta, who will in turn post on his government website, a report detailing their content-moderation practices, including what specifically they deem to be hate speech, racism, extremism, radicalization, disinformation, misinformation, harassment, and foreign political interference (or whether they refuse to define those terms), detailed statistics about how such speech has been moderated, and which of their policies are "intended to address th[ose] categories of content" (the "Terms of Service (or 'TOS') Report"), *see* §22677; and (iii) a penalty provision, whereby companies may be sued in court for, among other things, "[m]aterially omit[ting] or misrepresent[ing] required information in a" TOS Report and may be liable to pay up to $15,000 per violation per day, *see* §22678.[2] There is no guidance as to what constitutes a material omission or misrepresentation—a determination the statute leaves for the AG to define in his

[2] The court may, in its discretion, consider whether the company made a "reasonable, good faith attempt" to comply with the statute in "assessing the amount" of the penalty. §22678(a)(3).

sole discretion in deciding whether to pursue an enforcement action.

Below, X Corp. challenged the TOS Requirement (§22676), TOS Report (§22677), and penalty provisions (§22678) (as applied to both the TOS Requirement and TOS Report). This appeal challenges only the rulings on the TOS Report and penalty provisions as applied to it. AB 587 is attached hereto as an addendum.

AB 587's Purpose

The California legislature touts AB 587 as "a transparency measure" designed to "increase transparency around what terms of service social media companies are setting out" and "ensure[] those terms are abided by." *E.g.*, 4-ER-430. But a careful review of the law's purpose and likely effect—as evidenced by the text, legislative history, statements from AB 587's sponsors and supporters, and statements from AG Bonta in defending the law—demonstrates that AB 587 targets constitutionally protected speech based on content and viewpoint. The law's sponsors, and even AG Bonta himself, do nothing to hide this. Both the legislative history and AG Bonta's legal briefs openly concede that AB 587 seeks to apply "pressure" on social media companies into restricting constitutionally protected speech that the government finds objectionable:

i. The Assembly Committee on Judiciary Report about AB 587 states that, "if social media companies are forced to disclose what they do in this regard [i.e., how they moderate online content], it may ***pressure them*** to become better

corporate citizens by doing more ***to eliminate hate speech and disinformation.***" 4-ER-405.[3]

ii. AG Bonta has argued that "the Legislature also considered that, by requiring greater transparency about platforms' content-moderation rules and decisions, ***AB 587 may result in public pressure*** on social media companies to 'become better corporate citizens by doing more ***to eliminate hate speech and disinformation' on their platforms***." 5-ER-704–05.

iii. Within the Assembly Committee on Judiciary Report, lead bill author Jesse Gabriel stated that AB 587 is an ***"important first step"*** in ensuring that ***"social media companies [] moderate or remove hateful or incendiary content"*** on their platforms. He hoped that AB 587 will ***"pressure them" to "eliminate hate speech and disinformation."*** 4-ER-405.

iv. The Senate Judiciary Committee Hearing Report cites to comments from official bill sponsor, the Anti-Defamation League ("ADL"), that emphasized that the law will ***allow "policymakers [to] take meaningful action to decrease online hate and extremism."*** 4-ER-431.

v. On March 29, 2021, in a press release about AB 587, the ADL, an official

---

[3] Unless otherwise indicated, emphases in quotes are added and internal citations and quotations are omitted.

sponsor of AB 587, clarified that the intent of the law is to ***"improv[e]" the "enforcement of [social media companies' content-moderation] policies" or "provide enough evidence for legal action against them."*** 6-ER-930.

vi. On June 21, 2021, lead bill author Jesse Gabriel tweeted that AB 587 was going to ***"address . . . concerns that platforms aren't doing enough to stop the spread of misinformation and hate speech."*** 6-ER-932.

The legislative history also acknowledges that AB 587 focuses on categories of speech that are the most controversial to define and moderate. In fact, they are so controversial that almost any public commentary about how or whether to define or moderate them results in intense public criticism. According to the Assembly Report from the Committee on Privacy and Consumer Protection:

> [T]he companies responsible for managing social media platforms are faced with a complex dilemma regarding content-moderation, i.e., how the platforms determine what content warrants disciplinary action such as removal of the item or banning of the user. In broad terms, there is a general public consensus that certain types of content, such as child pornography, depictions of graphic violence, emotional abuse, and threats of physical harm are undesirable, and should be mitigated on these platforms to the extent possible. ***Many other categories of information, however, such as hate speech, racism, extremism, misinformation, political interference, and harassment [i.e., the categories that are the focus of AB 587], are far more difficult to reliably define, and assignment of their boundaries is often fraught with political bias. In such cases, both action and inaction by these companies seems to be equally maligned: too much moderation and accusations of censorship and suppressed speech arise; too little, and the platform risks fostering a toxic, sometimes dangerous community***.

4-ER-394. The unrebutted evidence in the record confirms that this accurately

describes X Corp.'s experience. Specifically, "[t]here is intense public debate and controversy about how to define the categories of content that should be limited on the social media platform and how to apply those categories to content on the social media platform. No matter what decisions are made, there are almost always large groups of people who disagree with them." 7-ER-1114. Content-moderation decisions are so controversial that X Corp. employees who publicly admit to taking part in that decision-making process are often harassed, abused, and attacked for their involvement. 6-ER-1088–91.

X Corp. Is Transparent About Its Content-Moderation Decisions

Significantly, prior to the passage of AB 587, X Corp. already provided its users with detailed information about (1) how it moderates content and (2) what kinds of content may lead to users being removed from the X platform. 7-ER-1114–17. There is no evidence that users of X—or of other social media platforms—feel they have insufficient information about these topics or need additional information to make informed choices about what platforms to use, such that a government mandate of the sort AB 587 imposes is necessary or even desirable.

Given the inherent controversy tied to the categories of content in §22677(a)(3), including that they are "difficult to reliably define" and "fraught with political bias," the task of moderating them, and doing so with the right balance, is extremely difficult. 7-ER-1113–14. X Corp. thus dedicates immense time, energy,

and financial and employee resources to these moderation efforts and to ensuring that they are fully accessible and understandable to X's users. 7-ER-1117. To that end, X's content-moderation policies are publicly-available on its website with clear explanations about how they are applied. 7-ER-1114–16. Those policies include, but are not limited, to X Corp.'s (1) Violent Speech, (2) Abuse and Harassment, (3) Hateful Conduct, (4) Violent and Hateful Entities, (5) Abusive Profile Information, (6) Crisis Misinformation, (7) Synthetic and Manipulated Media, and (8) Civic Integrity Policies. 4-ER-533–94. Some of these policies cover content that arguably falls within the controversial and difficult-to-define categories of speech that are the focus of AB 587.

AG Bonta Has Already Used AB 587 to Pressure X Corp. to Change Its Content-Moderation Policies

The TOS Reports, notwithstanding their seemingly innocuous title, are key to AB 587's amorphous and draconian penalty scheme, which allows the State to threaten to impose tremendous costs on social media companies if they do not moderate controversial content on their platforms in the State's preferred manner. The law affords the AG unfettered discretion in deciding what constitutes a "material[] omi[ssion] or misrepresent[ation]" of "required information" in the TOS Report. §22678(a)(2)(C). If the AG decides there is some reason to suspect that the compelled content-moderation statistics regarding hate speech, misinformation, and so forth are, in his view, incomplete or inaccurate, he is empowered to issue

compulsory demands for documents or testimony to investigate further. *See* Cal. Gov't Code §11180, *et seq*. And, even if no charges are ever filed, those compulsory requests can impose substantial costs on social media companies, so much so that the obvious response of many will be to avoid such inquiries altogether by modifying the content-moderation policies to appease the State. 7-ER-1112–13. As one leading commentator put it, "Through actual or threatened enforcement, regulators can influence what content Internet services publish—and punish Internet services for making editorial decisions the regulators disagree with." 6-ER-982.

Concerns about AB 587 being used in this way are not merely theoretical. Less than two months after AB 587's enactment, AG Bonta sent a threatening letter to the CEOs of X Corp. and other leading social media companies, reminding them of their "responsibility" to combat what AG Bonta described as the "dissemination of disinformation that interferes with our electoral system" and, at the same time, warning them that ***the "California Department of Justice will not hesitate to enforce" AB 587*** and other California laws. 6-ER-1070.

This threat of AB 587's enforcement was coupled with a series of carefully-worded demands from the AG in other sections of the letter. Specifically, his letter threatens enforcement while reminding its recipients of their "duty" to fight disinformation, urges them to "do more to rid [their] platforms of . . . dangerous disinformation," and implores them to use their "immense resources" to "take action

. . . against disinformation." 6-ER-1062–63, 1074, 1086. Any sophisticated reader of the letter would have reached the same conclusion as did Wifredo Fernandez, X Corp.'s Head of U.S. Governmental Affairs:

> [L]etters from Attorneys General, such as this one, that 'urge' companies to take action that the Attorney General claims they have a 'duty' or 'responsibility' to do, and, at the same time, threaten enforcement of certain specified laws, are a precursor to legal action taken by the Attorney General if the companies don't 'voluntarily' take the actions requested by the Attorney General.
>
> Based on my experience in governmental affairs and in dealing with numerous offices of Attorneys General across the country, I interpret Attorney General Bonta's letter as a thinly-veiled threat from the Attorney General to try to force X Corp. to limit specific speech—here, 'misinformation' or 'disinformation,' presumably as defined by Attorney General Bonta's Office–that Attorney General Bonta finds objectionable or face enforcement action.

6-ER-1064. AG Bonta did not and could not submit evidence to the district court suggesting anything to the contrary.

Of course, such threats from government actors chill constitutionally protected speech, even if there is no follow-through. AB 587 grants AG Bonta nearly unfettered discretion to determine if, in his view, X Corp. has made a "material[] omi[ssion] or misrepresent[ation]" in its TOS Report, and that view could be based on little more than his conviction that the statistics provided are

inconsistent with his assessment of the amount of disfavored content on the platform. §22678.

The end result is that AB 587 is not merely a "transparency measure." It provides the AG broad powers to pressure social media companies like X Corp. with threats of investigation and draconian fines if they fail to moderate content on their platforms in a manner that the State desires. As the legislative history makes plain, that is precisely what the State designed AB 587 to accomplish. And as the threatening letter from AG Bonta to X Corp.'s CEO makes plain, that is precisely how the State intends to use and is already using the law. 6-ER-1064.

Procedural History

On October 6, 2023, X Corp. moved for a preliminary injunction on the grounds that AB 587 violates its free speech rights under the U.S. and California Constitutions and that it is preempted by Section 230. 4-ER-595–97. X Corp. argued that the district court should apply strict scrutiny because, *inter alia*, AB 587 is a content- and viewpoint-based speech regulation and does not fall within the ambit of either *Zauderer* or *Central Hudson Gas & Elec. Corp.* v. *Pub. Serv. Comm'n of New York*, 447 U.S. 557 (1980), such that a lessened standard of review, appropriate only for commercial speech regulations, would apply. 5-ER-648–58.

On October 27, 2023, AG Bonta opposed X Corp.'s motion. 4-ER-333–86. Therein, AG Bonta conceded that "AB 587 is content based," 4-ER-373, but argued that *Zauderer* should apply. 4-ER-349–60.

At oral argument, the district court focused heavily on the appropriate standard of review—along the way suggesting, "I don't think this is commercial speech here," 2-ER-174—and ordered the parties to submit supplemental briefing on the appropriate standard of review. 2-ER-208–25.

On November 20 and November 28, 2023, AG Bonta and X Corp., respectively, submitted their supplemental briefing. 2-ER-124–57 (arguing that *Zauderer* applies); 2-ER-11–52 (arguing that strict scrutiny applies).

The Decision Below

On December 28, 2023, the district court issued an 8-page decision denying X Corp.'s preliminary injunction motion, holding that (1) *Zauderer* applied and X Corp. failed to establish a likelihood of success of a First Amendment violation and (2) Section 230 did not preempt AB 587. 1-ER-2–9.

Despite concluding that the TOS Reports "compelled by AB 587 do not so easily fit the traditional definition of commercial speech" because they "are not advertisements, and social media companies have no particular economic motivation to provide them," 1-ER-5, the district court "[n]evertheless" applied *Zauderer*—even though it only applies to compelled commercial speech—because it was

"following the lead of the Fifth and Eleventh Circuits" in the *NetChoice* cases. 1-ER-5. Those cases involved Florida and Texas statutes that, unlike AB 587, did not require social media companies to define (or decline to define) and provide statistics about specific categories of speech—let alone highly controversial categories of speech—provided by the State.

The district court held that *Zauderer* was satisfied because the TOS Reports were "purely factual" in that the required statistics "constitute objective data concerning the company's actions." 1-ER-5. The district court did not explain how a social media company's position on what constitutes hate speech, racism, harassment, and so forth is "purely factual" information. The district court also held that the TOS Reports were uncontroversial. On this point, the district court neither addressed the legislative history cited above, nor the inherently controversial nature of §22677(a)(3)'s topics, nor the unrebutted evidence that any statements made about how to define or moderate those topics result in public disapproval, nor that one of the law's stated goals was for the compelled disclosures to cause public pressure on the social media companies to moderate content to the State's liking. Rather, the district court held that "[t]he mere fact that the reports may be 'tied in some way to a controversial issue' does not make the reports themselves controversial." 1-ER-6. Finally, the district court concluded that the TOS Report requirement was not "unjustified or unduly burdensome within the context of First

Amendment law" because it did not "effectively rule[] out" the speech it accompanies. 1-ER-6.

The district court also held that Section 230 does not preempt AB 587 because it does not "provide for any potential liability stemming from a company's content-moderation activities *per se*." 1-ER-8. Although the law allows the State to pressure social media companies to limit or remove categories of content that the State disfavors by imposing significant investigative and enforcement costs on them and imposes liability on the companies for not regulating content in the statute's specified manner, the district court concluded that Section 230 does not preempt AB 587 because it does not "interfere with companies' ability to 'self-regulate offensive third party content without fear of liability." 1-ER-8.

On January 12, 2023, X Corp. timely filed notice of its appeal. 6-ER-1094–95.

SUMMARY OF THE ARGUMENT

The district court erred by finding that X Corp. did not establish a likelihood of success on the merits and denying its preliminary injunction motion.

I. AB 587's TOS Report requirement violates the First Amendment to the U.S. Constitution and Article I, Section 2, of the California Constitution.[4]

A. The district court erred by applying *Zauderer* to the TOS Report because (1) the TOS Report's compelled disclosures are not commercial speech, which the district court itself appeared to accept; (2) the TOS Report's compelled disclosures are not purely factual, uncontroversial information about the terms under which X's services will be provided; (3) the TOS Report does not further a substantial government interest; and (4) even if the speech compelled by the TOS Report were commercial, it is inextricably intertwined with core political speech and therefore must receive full First Amendment protection.

B. Strict scrutiny applies to AB 587 because (1) AB 587 is a content-based, viewpoint-discriminatory regulation of speech, and no exception applies to the longstanding rule that such regulations trigger strict scrutiny. *See Nat'l Inst. of Fam. & Life Advocs.* v. *Becerra* ("*NIFLA*"), 585 U.S. 755, 767 (2018) ("This Court's precedents do not permit governments to impose content-based restrictions on speech without persuasive evidence . . . of a long (if heretofore unrecognized)

---

[4] AB 587 violates Article I, Section 2, of the California Constitution for the same reasons that it violates the First Amendment. *See, e.g.*, *City of Montebello* v. *Vasquez*, 1 Cal. 5th 409, 421 n.11 (2016) (California's free speech protections are broader than those provided by the First Amendment); *Delano Farms Co.* v. *Cal. Table Grape Com.*, 4 Cal. 5th 1204, 1221 (2018).

tradition to that effect."); (2) AG Bonta has already used AB 587's statutory scheme to pressure X Corp. to moderate content that the State disfavors, thereby impermissibly interfering with X Corp.'s constitutionally protected editorial judgment over how to moderate user content on its platform; and (3) AB 587 is a regulation of "speech about speech" that not only infringes *X Corp.'s* constitutionally protected ability to moderate content on its platform as it sees fit, but also disrupts the ability of *X users* to view and disseminate content on the platform.

C. AB 587 fails under any level of scrutiny. The TOS Report fails strict scrutiny because its requirements are not narrowly tailored to a compelling government interest, and there are less-restrictive alternatives that would serve the State's stated goal of providing consumers with information to make informed choices about what social media platforms to use. The TOS Report also fails intermediate scrutiny under *Central Hudson* because the State's purported interest is not substantial and the TOS Report's compelled disclosure requirements are more extensive than necessary to further that interest. Moreover, the State failed to demonstrate that the TOS Report alleviates any harms that—at least as to X Corp.—are real, or that the statute would alleviate them to a material degree. Finally, even if *Zauderer* applied (and it does not), the TOS Report still does not withstand constitutional review because its disclosure requirements will chill not just X Corp.'s

protected speech, but also that of X users. Nor is the TOS Report reasonably related to a substantial government interest, which renders the law unconstitutional under *Zauderer* as well.

II. The district court also erred in finding that Section 230 does not preempt AB 587 because any penalization of X Corp. for moderating content covered by §230(c)(2) without making the TOS Report's compelled disclosures to the State's satisfaction would not be "liability stemming from a company's content-moderation activities *per se*." 1-ER-8. Section 230 immunity precludes the liabilities AG Bonta can impose under AB 587 by issuing costly discovery demands and threatening and bringing enforcement actions intended to pressure social media companies to change their content-moderation policies. It also precludes the AG from requiring content-moderation practices that have the purpose and likely effect of pressuring companies to change their content-moderation policies.

STANDARD OF REVIEW

This Court reviews *de novo* decisions denying a preliminary injunction on the basis that there was no likelihood of success on the merits. *Associated Gen. Contractors of Am.* v. *Metro. Water Dist. of S. Cal.*, 159 F.3d 1178, 1180 (9th Cir. 1998); *N.D. ex rel. parents acting as guardians ad litem* v. *Hawaii Dep't of Educ.*, 600 F.3d 1104, 1111 (9th Cir. 2010) ("[W]e 'determine *de novo* whether the trial court identified the correct legal rule to apply to the relief requested.'").

Accordingly, reversal is warranted if the district court "based its decision on an erroneous legal standard." *Taylor* v. *Westly*, 488 F.3d 1197, 1199 (9th Cir. 2007).

To succeed on a preliminary injunction motion, a plaintiff must establish that (1) it is likely to succeed on the merits; (2) it is likely to suffer irreparable harm absent preliminary relief; (3) the balance of equities tips in its favor; and (4) an injunction is in the public interest. *Winter* v. *Nat. Res. Def. Council, Inc.*, 555 U.S. 7, 20 (2008).

ARGUMENT

I. AB 587 VIOLATES THE FIRST AMENDMENT AND ARTICLE 1, SECTION 2, OF THE CALIFORNIA CONSTITUTION

A. Neither *Zauderer* Nor *Central Hudson* Applies To AB 587

1. AB 587 Compels Non-Commercial Speech

The district court erred because, despite effectively concluding that the TOS Report's compelled disclosures are not commercial speech, *see, e.g.*, 1-ER-5 (the disclosures "do not so easily fit the traditional definition of commercial speech"), it nevertheless followed the Eleventh and Fifth Circuits' application of *Zauderer* to other, readily distinguishable state disclosure statutes in the *NetChoice* cases, 1-ER-5. The district court was correct to suggest that the compelled speech here is not commercial and therefore erred in applying *Zauderer*.

Zauderer review applies only to commercial speech. *See, e.g.*, *Nat'l Ass'n of Wheat Growers* v. *Bonta*, 85 F.4th 1263, 1266 (9th Cir. 2023) ("[*Zauderer*] applies

to compelled ***commercial*** speech that is 'purely factual and uncontroversial.'"). As both the Supreme Court and this Court have made clear, the correct test for identifying commercial speech is whether it "does no more than propose a commercial transaction." *IMDb.com Inc.* v. *Becerra*, 962 F.3d 1111, 1122 (9th Cir. 2020) (quoting *Va. State Bd. of Pharmacy* v. *Va. Citizens Consumer Council, Inc.*, 425 U.S. 748, 762 (1976)); *Hunt* v. *City of Los Angeles*, 638 F.3d 703, 715 (9th Cir. 2011) (same) (quoting *United States* v. *United Foods*, Inc., 533 U.S. 405, 409 (2001)).[5] AB 587's compelled disclosures plainly fail this test. The disclosures have various purposes. For instance, they convey X Corp.'s policy views on what constitutes hate speech, racism, misinformation, and so forth (or X Corp.'s view that it should not define or police those controversial categories); and they convey information about how X Corp. has applied its policies, and, in so doing, convey information about what kind of speech X Corp. believes should be tolerated on X. They do not, however, propose a commercial transaction in any way.

This Court has also suggested that in cases "[w]here the facts present a close question" as to whether the speech "does no more than propose a commercial

---

[5] At least four additional Supreme Court cases have defined commercial speech as that which "does no more than propose a commercial transaction." *Bolger* v. *Youngs Drug Prod. Corp.*, 463 U.S. 60, 66 (1983); *City of Cincinnati* v. *Discovery Network, Inc.*, 507 U.S. 410, 422 (1993); *Bd. of Tr.* v. *Fox*, 492 U.S. 469, 482 (1989); *Central Hudson*, 447 U.S. at 562.

transaction," courts will look to the factors listed in *Bolger*, 463 U.S. at 66–67, to determine if the speech is commercial—i.e., whether (1) "the speech is an advertisement," (2) "the speech refers to a particular product," and (3) "the speaker has an economic motivation." *Hunt,* 638 F.3d at 715. If all three factors are present, then a strong case can be made that the speech is commercial. *Id.*

As the district court correctly suggested, however, AB 587's compelled disclosures fail to satisfy at least two of these factors because they "are not advertisements, and social media companies have no particular economic motivation to provide them." 1-ER-5. Moreover, as the district court suggested at oral argument, the context of the disclosures here further undermines any suggestion that they are commercial. The TOS Report is submitted to the AG and published on a government website, not provided to consumers at the point of sale. *See* 2-ER-202–03 (THE COURT: "a disclosure to the Government is [not] a disclosure to the public" that would "bring[] [the disclosures] into the realm of commercial speech," since the "Attorney General is not the consumer.").

The district court's decision to apply *Zauderer*—primarily because the Fifth and Eleventh Circuits did so in the *NetChoice* cases to statutes that the district court mistakenly believed to be "strikingly similar" to AB 587—was clear error. The district court notably did not conclude that the rationale of the Fifth and Eleventh Circuits for applying *Zauderer* was persuasive. It merely cited their holding and

said, without explanation, that it was "[f]ollowing the[ir] lead." 1-ER-5.

But the disclosure provisions of the laws in the *NetChoice* cases (Florida SB 7072 and Texas HB 20) are actually not "strikingly similar" to AB 587; they fundamentally differ in legally significant ways. Unlike those statutes, the TOS Report compels disclosures that single out specific categories of constitutionally protected content. *See, e.g.*, §22677(a)(4)(A) (forcing social media company to disclose a "detailed description" of whether its "content-moderation practices" are "intended to address" hate speech, racism, and so forth). In contrast, none of Florida SB 7072's or Texas HB 20's required disclosures hinged on the type of content at issue. *NetChoice (Fla.)*, 34 F.4th at 1206–07 (requiring platforms to disclose information about their content-moderation "standards," "rule changes," "view counts for posts," "free advertising for political candidates," and "explanations" for content-moderation decisions); *NetChoice (Tex.)*, 49 F.4th at 446 (requiring platforms to disclose "how they moderate and promote content," provide "high-level statistics related to [those] efforts," and maintain "a complaint-and-appeal system" for content-moderation disputes). Not only do the TOS Report's compelled disclosures target particular categories of content, but, as the legislative history and undisputed evidence in the record make plain, those categories are "among the most difficult to define and most controversial to apply," 7-ER-1114, and the law seeks to pressure the social media companies to limit or eliminate content in those

controversial categories. 6-ER-1064; 4-ER-405; 5-ER-738; 5-ER-704. Again, neither is true of the Florida or Texas laws.

The TOS Report's disclosure requirements are far more analogous to those in the New York law in *Volokh* v. *James*, which "regulate[d] speech based on its content" by forcing social media networks to publish certain policies and statistics about moderation of "hateful conduct." 656 F.Supp.3d 431, 443–44 (S.D.N.Y. 2023). While the law in *Volokh* defined "hateful conduct," *id*. at 437, the law, like AB 587, did not require the social media companies to adopt any particular policy toward hateful conduct or to take action against it. *See id*. at 442 ("[T]he Hateful Conduct Law ostensibly does not dictate what a social media website's response to a complaint must be and does not even require that the networks respond to any complaints or take down offensive material"). Like AB 587, it merely sought to pressure social media companies to do more to limit the spread of hateful conduct by forcing them to disclose their policies (or lack thereof) for addressing such conduct. *See* 4-ER-405; 5-ER-738; 5-ER-704.

Given the New York statute's focus on "hateful conduct" and the pressures it created to change content-moderation practices about such speech, the court applied strict scrutiny and struck it down. *See Volokh*, 656 F.Supp.3d at 443. The court held that the compelled disclosures did "not constitute commercial speech" because the "law's requirement that Plaintiffs publish their policies explaining how they intend

to respond to hateful content on their websites does not simply 'propose a commercial transaction.' Nor is the policy requirement 'related solely to the economic interests of the speaker and its audience.' Rather, the policy requirement compels a social media network to speak about the range of protected speech it will allow its users to engage (or not engage) in." *Id.*

As legal commentators have recognized, this focus on categories of content selected by the State distinguishes the statute in *Volokh* and AB 587 from the statutes in the *NetChoice* cases and requires application of strict scrutiny, even if the speech category-neutral laws in *NetChoice* are subject to a different standard of review:

> Platform transparency mandates like the ones in New York [at issue in *Volokh*] and California [i.e., AB 587's TOS Report] create burdens and enforcement risks that, like those in *NetChoice*, may affect the platforms' actual editorial functions. But they do so only for particular kinds of editorial policies. The result, recognized in [*Volokh*] . . . is a content-based burden on editorial processes [that warrants application of strict scrutiny].

6-ER-1024. The district court erred in not recognizing this fundamental distinction and blindly "[f]ollowing the lead" of the courts in the *NetChoice* cases.

2. AB 587's Compelled Disclosures Are Not Purely Factual, Uncontroversial Information About The Terms Under Which X Corp.'s Services Will Be Available

The district court also erred by applying *Zauderer* because AB 587's compelled disclosures are (1) not purely factual, (2) highly controversial, and (3) not about the terms under which X Corp. makes its services available.

First, the district court erred by finding that the TOS Report's compelled disclosures are "purely factual" because "they constitute objective data concerning the company's actions." 1-ER-5. The TOS Report's compelled disclosures are not "purely factual," because expressing a position on what constitutes hate speech, racism, misinformation, disinformation, harassment, extremism, radicalism, and foreign political interference—or publicly refusing to do so—is based on highly subjective and value-based judgments. Moreover, they are bound up in a platform's own editorial opinions and policy discretion.

While the Supreme Court has not clarified what constitutes "purely factual" information for *Zauderer* purposes, the qualifier "purely" strongly suggests that it was not intended to apply to the types of disclosures about content-moderation that are governed by AB 587. Is a platform's policy that it is not harassment to refer to a transgender woman as a man "***purely*** factual"? Is a platform's policy that the phrase "From the River to the Sea" constitutes hate speech when made in reference to the Israeli-Palestinian conflict, but not in other contexts, "***purely*** factual"? Is a platform's disclosure that reports about Hunter Biden's laptop constitute foreign political interference "***purely*** factual"? Of course not. As one legal commentator has stated:

> [I]t might be a pure fact whether the publisher has an editorial policy, but it would not be a pure fact to disclose the policy's details, which remain subject to the publisher's editorial discretion. . . . Similarly, consider California's requirement that an online publisher disclose the number of items it has

> 'actioned.' This disclosure is impossible because it assumes there are only two outcomes for any content item: actioned or not. In reality, every item is prioritized or deprioritized relative to other items. Furthermore, if [social media companies] personalize content ordering, items may be 'actioned' for some readers and not actioned for others. This complexity and ambiguity makes it impossible to characterize 'actioned' as a 'purely factual' disclosure.

3-ER-323. Here, §22677(a)(4)(A), for example, requires X Corp. to provide a "detailed description of content-moderation practices used by the social media company for [its] platform, including, but not limited to, . . . [a]ny existing policies *intended to address* the categories of content described in paragraph (3)." It is difficult to see how providing a "detailed description" of which of its policies are "intended to address" §22677(a)(3)'s controversial categories can be a disclosure of purely factual information. Whether a company has adopted a policy that is "intended to address" difficult-to-define categories of constitutionally protected speech requires the exercise of judgment and expression of opinions that could be the subject of disagreement among reasonable individuals. It is certainly not "purely" factual.

Even the disclosure of statistics such as the total number of items flagged by the social media company as falling within the various categories of content, *see* §22677(a)(5)(A), is not purely factual, because placing a user's post in a particular category is itself an act of judgment. *See Book People, Inc.* v. *Wong*, 91 F.4th 318, 340 (5th Cir. 2024) ("The statute requires vendors to undertake contextual analyses, weighing and balancing many factors to determine a rating for each book. Balancing

a myriad of factors that depend on community standards is anything but the mere disclosure of factual information."). Similarly, the act of explaining why posts were flagged—e.g., by "disaggregat[ing]" those statistics by content category, *see* §22677(a)(5)(B)(i)—is not purely factual either.

Second, the district court erred in finding that the TOS Report's compelled disclosures are uncontroversial. 1-ER-5–6. Relying on this Court's decision in *CTIA - The Wireless Ass'n* v. *City of Berkeley* ("*CTIA II*"), 928 F.3d 832, 845 (9th Cir. 2019), the district court held that, "[t]he mere fact that the reports may be 'tied in some way to a controversial issue' does not make the reports themselves controversial." 1-ER-6.

But the disclosures here are not merely "tied in some way to a controversial issue" the way that the cell phone warnings in *CTIA II* were. AB 587 compels social media platforms to disclose their policy opinions about controversial issues, such as what constitutes hate speech, or publicly refuse to do so. The controversial issue is the crux of the disclosure. Indeed, as the legislative history and statements from AG Bonta make clear, AB 587 is *designed* to *generate* public controversy about the actions of the social media companies by framing a divisive debate in such a way that will pressure the social media companies to change their content-moderation policies to align with those desired by the State. That is why the statute focuses on what the legislative history concedes are the most controversial and politically-

charged categories of content that, *no matter how they are moderated*, will invariably leave some set of users dissatisfied. *See, e.g.*, 5-ER-738; 5-ER-746.

Put another way, the assertion that the TOS Report's disclosures are merely tied in some way to a controversial issue cannot be reconciled with AG Bonta's own statements that one of the law's purposes is to apply "public pressure" that will lead the companies to change their content-moderation policies. 5-ER-704–05. The very notion that the disclosures will result in "public pressure" on the social media companies to change their policies presupposes that the disclosures are so controversial that they will result in the public demanding such changes.

The compelled disclosures are also controversial because they force X Corp. to "take sides in a heated political controversy." *CTIA II*, 928 F.3d at 848. The act of defining (or choosing not to define) §22677(a)(3)'s categories is inherently controversial. And providing statistics about moderation of these topics is also controversial because there is widespread disagreement on how much (or how little) social media companies should do to limit such speech on their platforms.[6]

---

[6] That the statute does not explicitly mandate that X Corp. adopt the State's views does not make the disclosure uncontroversial. For example, forcing a person to express their views on abortion would be controversial, even if they were not required to express a particular viewpoint. Similarly, forced public disclosure of X Corp.'s own position on highly contentious topics in response to questions formulated by the State is controversial in its own right.

The compelled disclosures are also controversial because they force X Corp. to "convey a message fundamentally at odds with its mission" and "business." *Wheat Growers*, 85 F.4th at 1277–78. By forcing social media companies to take a public stance on whether to define and regulate the statute's categories, AB 587 allows the State to frame the debate about content moderation—by suggesting categories of speech that the State seems to think should be disfavored or censored—rather than allowing the companies to frame the debate themselves.

This Court's recent decision in *Wheat Growers*, moreover, confirms that there is also a subjective component to determining whether compelled disclosures are controversial under *Zauderer*. *Id.* at 1277 ("*NIFLA* tells us that the topic of the disclosure and its effect on the speaker is probative of determining whether something is subjectively controversial."). The TOS Report's compelled disclosures flunk the subjective part of the test quite dramatically. As the legislative history makes abundantly clear, the categories of speech that AB 587 asks social media companies to define—or announce that they have chosen not to define—are those having no "general public consensus" about how to define or moderate them because their boundaries are "fraught with political bias" and are "difficult to reliably define." 5-ER-746. And the detailed information that the companies are compelled to provide about their content-moderation decisions concern the most controversial

decisions that will lead them to be "equally maligned" by members of the public, whether they restrict such content or not. 5-ER-746.

Finally, the district court should not have applied *Zauderer* because the TOS Report's compelled disclosures are not "purely factual and uncontroversial information *about the terms under which [X Corp.'s] services will be available*." *NIFLA*, 585 U.S. at 768 (quoting *Zauderer*, 471 U.S. at 651). Statistics about how X Corp. moderated content on X during the previous quarters or how X Corp. defined a particular content category within §22677(a)(3) during that period, are backward-looking statements that do not provide information about the terms under which the X platform will be made available. The TOS Report's disclosures are also not an explication of the services being provided—they are, instead, a compilation of opinion-based judgments that go well beyond mere descriptive "terms."

3. AB 587 Does Not Further A Substantial Government Interest

The district court also erred by finding that the government's interest in "requiring social media companies to be transparent about their content-moderation policies and practices so that consumers can make informed decisions about where they consume and disseminate news and information" is "substantial" or "more than trivial." 1-ER-6–7 (quoting *CTIA II*, 928 F.3d at 844). To date, the only interests that have been recognized by the Supreme Court or this Court as sufficiently substantial to trigger *Zauderer* review are prevention of consumer deception and, in

some instances, "furthering public health and safety." 2-ER-38. While this Court made clear in *CTIA II* that it was not "foreclos[ing] that other substantial interests in other cases may suffice as well,"[7] it also clarified that, to qualify, the "interest at stake must be more than the satisfaction of mere 'consumer curiosity.'" *CTIA II*, 928 F.3d at 843–44 (quoting *Nat'l Elec. Mfrs. Ass'n* v. *Sorrell*, 272 F.3d 104, 115 n.6 (2d Cir. 2001), which cited to *Int'l Dairy Foods Ass'n* v. *Amestoy*, 92 F.3d 67 (2d Cir. 1996) (Vermont law likely violated First Amendment by requiring labeling of milk from cows given rBGH, even though the milk from such cows was indistinguishable from milk from untreated cows)). Thus, under *CTIA II*, "transparency" is not a substantial interest in itself, unless it serves some other important goal. That is precisely why this Court held that the "interest at stake must be more than the satisfaction of mere 'consumer curiosity.'" *CITA II*, 928 F.3d at 844.

The TOS Report does not further a sufficiently substantial governmental interest to trigger *Zauderer*. There is simply no proof in the record—and it was AG Bonta's burden to provide it—that the law prevents consumer deception or furthers

---

[7] Although this Court has determined that *Zauderer* applies outside the context of consumer deception, *CTIA II*, 928 F.3d at 844, X Corp. expressly preserves the right to argue otherwise on appeal, *see, e.g.*, *Milavetz, Gallop & Milavetz, P.A.* v. *United States*, 559 U.S. 229, 250, 253 (2010) (*Zauderer* applies to disclosures intended to prevent consumer deception).

a "public health and safety" interest of the type that would satisfy the "substantial interest" threshold. AG Bonta provided no evidence that (1) there was a need for more transparency concerning the content-moderation practices of social media companies or (2) the disclosures in the TOS Report would provide information that would allow consumers to make "informed choices" about what social media platforms to use. Put another way, there is no evidence that AB 587 addresses a real problem and that consumers actually need the information provided by the statute (on top of the already-public content-moderation policies provided by social media companies) to make informed decisions about what platforms to use.

The information compelled by AB 587, while burdensome to compile, does not even provide enough detail to adequately assess how often platforms are removing particular kinds of content. For example, AB 587 does not compel information that would reveal how many posts are on the social media platform overall, so the Report would not reveal the percentage of posts that are actioned or how often speech that arguably falls within these categories of content was *not* actioned. Nor do the statistics reveal anything that would tell a potential user of social media *why* the statistics for moderation of certain categories of speech are what they are. Are low numbers of actioned content because such content rarely appears or because the definition of the content is being applied narrowly? While "transparency" sounds like a noble interest in the abstract, there is simply no proof

that the statistics here provide anything meaningful to potential social media users. Perhaps that is why, in the real world, it is difficult to imagine any potential user actually going to the AG's website and using the TOS Report to determine what platform to use.

Indeed, as the district court recognized at oral argument, because AB 587 allows social media companies to define the categories of content however they want (or not at all), companies' statistics are not readily comparable, so it is not clear how the statute provides consumers with statistics that are useful to make informed choices about anything. *See* 2-ER-210–11. The statute also does not require social media companies to make any statistical disclosures at all if their categories of content moderation differ from those listed in the statute. 2-ER-210–11; §22677. Accordingly, it is not clear that the statute will do anything other than incentivize social media companies to use content-moderation categories different from those in the statute—which would further no substantial government interest at all, and indeed would reduce transparency. *See* 2-ER-195.

In the end, there is no evidentiary basis for concluding that AB 587's compelled disclosures will actually help consumers "make informed decisions about where they consume and disseminate news and information," 1-ER-6, and what remains is a "purely hypothetical" harm and desire to satisfy "mere 'consumer

curiosity,'" *CTIA II*, 928 F.3d at 844, which are insufficient to trigger *Zauderer* review.

4. AB 587 Compels Speech That Is Inextricably Intertwined With Otherwise Fully Protected Speech

Even if the speech here were commercial (and it is not), *Zauderer* would still not apply because the speech is, at the very least, "inextricably intertwined with otherwise fully protected speech," such that it does not "retain its commercial character" and must receive full protection under the First Amendment. *Hunt*, 638 F.3d at 715 (quoting *Riley* v. *Nat'l Fed'n of the Blind of N. Carolina, Inc.*, 487 U.S. 781, 796 (1988)).

The TOS Report's compelled disclosures provide information about the content-moderation decisions of social media companies. Those include whether (and if so, how) to define categories of constitutionally protected speech that are "fraught with political bias" and how those categories of content were moderated. 5-ER-746. They clearly and purposely require the platforms to reveal their positions on hotly debated and controversial political topics. Take for example, the requirement to disclose whether the social media company believes that "hate speech" or "racism" are categories of speech that should be disfavored on its platform and, if so, how those controversial categories should be defined. The answers to those questions are, by their nature, political. They reveal a vision of

how debate should be structured on social media platforms to promote the public good and how difficult-to-define categories of speech should be defined.

Likewise, take AB 587's requirement to provide a "detailed description" of "[a]ny existing policies intended to address" the categories in the statute. *See* §22677(a)(4)(A). If, in response to this requirement, X Corp. included in its TOS Report a "detailed description" of its Hateful Conduct Policy—which states, among other things, that "X's mission is to give everyone the power to create and share ideas and information, and to express their opinions and beliefs without barriers," 4-ER-549—there is no way that X Corp. could separate this so-called 'commercial information' (if it is that) from the "political or social issues" with which it is "inextricably intertwined." *Gaudiya Vaishnava Soc.* v. *City & Cnty. of San Francisco*, 952 F.2d 1059, 1063–64 (9th Cir. 1990).

The same is true of the mandate to provide statistics flowing from those decisions. For instance, the TOS Report mandates that X Corp. provide the total number of flagged and actioned items "belonging to" the categories of content in §22677(a)(3). *See* §22677(a)(5)(A). But, similarly, X Corp. cannot provide that information without revealing its views on those core political topics. Thus, even assuming *arguendo* that the compelled disclosures are commercial speech, they still could not be subject to lessened First Amendment protection under *Zauderer*,

because it would be impossible to separate the purportedly commercial components of X Corp.'s speech from the core political components.

B. Strict Scrutiny Applies To AB 587

1. AB 587 Is A Content- And Viewpoint-Based Regulation Of Speech

It is plain as day that AB 587 targets particular speech based on its content. It singles out, for differential treatment, moderation of specific categories of speech, and creates additional burdens on social media companies moderating such content. AG Bonta has conceded that "AB 587 is content based," 4-ER-373, and at oral argument the district court appeared to agree, 2-ER-180 ("I think I said [AB 587] was content-based."). Under longstanding Supreme Court precedent, such "content-based laws—those that target speech based on its communicative content—are presumptively unconstitutional" and may "stand only if they survive strict scrutiny." *Reed*, 576 U.S. at 163, 171.

Not only does AB 587 target the particular categories of content set forth in §22677(a)(3), it discriminates against particular viewpoints within those categories, making it subject to strict scrutiny. *Rosenberger* v. *Rector & Visitors of Univ. of Va.*, 515 U.S. 819, 830–31 (1995). This is clear for at least the following two reasons. First, Sections 22677(a)(4) and (5) of the law, on their face, impose additional burdens and reporting requirements on X Corp. if it regulates certain categories of constitutionally protected speech. The very selection of these

categories strongly suggests that the State's view is that they are types of speech that ***should*** be subject to content-moderation decisions. Second, AB 587's legislative record, statements made by AG Bonta in defending and preparing to enforce the law, and the record in this case all make clear that the State's focus on §22677(a)(3)'s categories of content is not neutral. AB 587's main purpose is to "pressure" X Corp. to "eliminate" these categories from its platform because the State views them as problematic. *See, e.g.*, 5-ER-738; 5-ER-704–05; 6-ER-1070, 1073.

Whether the law is content-based, viewpoint-based, or both, it is clear that "[f]ormal legislative findings accompanying" AB 587 make clear its illicit "purpose and practical effect," *Sorrell* v. *IMS Health, Inc.*, 564 U.S. 552, 565 (2011), which is to regulate speech on X "based on 'the topic discussed or the idea or message expressed.'" *City of Austin, Texas* v. *Reagan Nat'l Advert. of Austin, LLC*, 596 U.S. 61, 73–74 (2022). This the First Amendment does not permit, absent a compelling state interest and means that are narrowly tailored to achieve that interest, since "subtler forms of discrimination that achieve identical results based on function or purpose" do not escape strict scrutiny. *Id*. at 74; *Reed*, 576 U.S. at 166 ("[S]trict scrutiny applies either when a law is content based on its face ***or*** when the purpose and justification for the law are content based[.]"); *Animal Legal Def. Fund* v. *Wasden*, 878 F.3d 1184, 1204 (9th Cir. 2018) (same).

2. AB 587 Impermissibly Interferes With X Corp.'s Constitutionally Protected Editorial Judgment

Strict scrutiny should also apply because (1) X Corp.'s right to moderate content on its platform as it sees fit is constitutionally protected, (2) compelled speech regimes, such as AB 587, interfere with that constitutionally protected editorial judgment, and (3) AG Bonta has ***already*** used AB 587's statutory scheme to try to coerce X Corp. into moderating particular content.

First, the editorial judgment that X Corp. enjoys over the X platform is First Amendment-protected, the same as if it were "a newspaper or a news network." *O'Handley* v. *Padilla*, 579 F.Supp.3d 1163, 1186–87 (N.D. Cal. 2022), *aff'd sub nom. on other grounds*, *O'Handley* v. *Weber*, 62 F.4th 1145 (9th Cir. 2023) ("Like a newspaper or a news network, [X Corp.] makes decisions about what content to include, exclude, moderate, filter, label, restrict, or promote, and those decisions are protected by the First Amendment."); *NetChoice (Fla.)*, 34 F.4th at 1213.

That is because the Supreme Court has made clear that the First Amendment protects "both the right to speak freely and the right to refrain from speaking at all." *Wooley* v. *Maynard*, 430 U.S. 705, 714 (1977). The right to choose whether and how to tailor one's message is not "restricted to the press," and instead applies equally to "business corporations generally" and "ordinary people engaged in unsophisticated expression as well as professional publishers." *Hurley* v. *Irish-Am. Gay, Lesbian & Bisexual Grp. of Boston*, 515 U.S. 557, 574 (1995). If there were a

law, for instance, compelling newspapers to provide detailed disclosures about their criteria for publication and statistics about the bases for decisions about whether to publish letters to the editor or op-eds, it would undoubtedly interfere with the constitutionally protected editorial judgment of newspapers.

That X is a social news application does not change the analysis. The Supreme Court has struck down on First Amendment grounds governmental attempts to inhibit "private entities' rights to exercise editorial control over speech and speakers on their properties or platforms." *Manhattan Cmty. Access Corp.* v. *Halleck*, 139 S. Ct. 1921, 1932 (2019). In *Halleck*, the Court held that a cable channel's decisions about what third-party content to allow on its channel was subject to full First Amendment protection. *Id.* A social media company's editorial judgment about what user content to allow on its platform is no different.

Second, compelled speech regimes like AB 587 that force social media companies to make disclosures regarding their content-moderation policies and decision-making inherently interfere with those companies' constitutionally protected editorial judgments. That interference, moreover, occurs well before the issue would actually reach a court—the threat of enforcement alone is enough. As one leading commentator has noted:

> These disclosure obligations send the message to publishers that regulators will be scrutinizing those disclosures. That message is coupled with an impossible-to-ignore threat that regulators will pursue the publisher if the disclosures do not satisfy the regulators' normative objectives[.] . . . ***Knowing***

> ***this risk, publishers will distort their editorial decisions so their disclosures placate regulators and mitigate the risk of future investigations or enforcement***. Inevitably, compelled editorial transparency changes the publisher's constitutionally protected editorial decision-making and affects constitutionally protected speech—exactly like an outright speech restriction would. ***It does not matter if an investigation or enforcement would be unconstitutional once pursued because the speech harms occur well before then***.

3-ER-318–19.

AB 587 has already resulted in such threats. AG Bonta has ***already*** threatened to enforce AB 587 against X Corp. if it does not do more to moderate content in ways the State desires. To that end, X Corp. submitted an affidavit from Wifredo Fernandez (its Head of U.S. Government Affairs) that attached a November 3, 2022 letter from AG Bonta to the CEOs of major social media companies (including X Corp.), wherein the AG proclaimed he would "not hesitate to enforce" AB 587 against the companies if they do not do more to remove or limit on their sites what AG Bonta called "disinformation" and "misinformation." 6-ER-1070. The letter repeatedly emphasized that, in AG Bonta's view, the companies had a "duty," "responsibility," and "obligation" to do more to combat these kinds of speech, and, in the same proverbial breath, reminded the companies that the AG would not hesitate to enforce AB 587. 6-ER-1070. Mr. Fernandez concluded that "letters from Attorneys General, such as this one, that 'urge' companies to take action that the Attorney General claims they have a 'duty' or 'responsibility' to do, and, at the same time, threaten enforcement of certain specified laws, are a precursor to legal action

taken by the Attorney General if the companies don't 'voluntarily' take the actions requested by the Attorney General." 6-ER-1064.

The effect of such a threat is magnified by the fact that, under AB 587, AG Bonta maintains nearly unfettered discretion to determine what constitutes a "material[] omi[ssion] or misrepresent[ation]" in violation of the statute, §22678(a)(2)(C), and whether to bring "an action pursuant to" AB 587 based on any such alleged violation, *see* §22678(c). Although it is "the court" that will ultimately "assess[] the amount of a civil penalty" for any such violation, §22678(a)(3), as explained, the pressure applied to X Corp. through threats of enforcement of AB 587's statutory scheme begins well before this issue would reach a court.

The impact of such threats is further magnified by the "broad pre-litigation powers" afforded to AG Bonta under California law, including but not limited to "issu[ing] subpoenas" for the "production of . . . documents," the "attendance of witnesses," and "testimony," Cal. Gov't Code §11181(a).

When these three factors are combined—i.e., AG Bonta's (1) threat to enforce AB 587 in a letter demanding changes to content-moderation policies; (2) broad discretion in determining whether TOS Reports contain a "material omission or misrepresentation"; and (3) ability to impose substantial litigation costs on social media companies by issuing document demands and compelling testimony if, in his sole discretion, he believes there is a potential violation—the chilling effect on

speech is substantial and more than sufficient to support injunctive relief, even without a pending prosecution. *See, e.g.*, *MedImmune, Inc.* v. *Genentech, Inc.*, 549 U.S. 118, 128–29 (2007) ("[W]here threatened action by *government* is concerned, we do not require a plaintiff to expose himself to liability before bringing suit to challenge the basis for the threat—for example, the constitutionality of a law threatened to be enforced.") (emphasis in original).

AB 587 is, accordingly, an impermissible government intrusion on free speech that "subjects [X Corp.'s] editorial process to private or official examination merely to satisfy curiosity or to serve some general end such as the public interest," which does "not survive constitutional scrutiny as the First Amendment is presently construed." *Herbert* v. *Lando*, 441 U.S. 153, 174 (1979); *Miami Herald Pub. Co.* v. *Tornillo*, 418 U.S. 241, 258 (1974) (government regulation of newspaper's editorial judgment violates First Amendment). Accordingly, AB 587 triggers strict scrutiny, under which it should be struck down.

The Fourth Circuit's decision in *Washington Post* v. *McManus*, 944 F.3d 506 (4th Cir. 2019), is directly on point. There, the Fourth Circuit struck down, on First Amendment grounds, a Maryland law requiring online platforms, as part of an effort to address foreign interference in U.S. elections, to disclose data about the identity of purchasers of advertising and make it available to the government for inspection upon request. *Id.* at 511–12. Finding the law to be a "compendium of traditional

First Amendment infirmities" that "brings the state into an unhealthy entanglement with news outlets," the court emphasized that the law's "specter of a broad inspection authority, coupled with an expanded disclosure obligation," had the potential to "chill speech and is a form of state power the Supreme Court would not countenance." *Id.* at 513, 518–19. The same is true of the TOS Report here.

3. AB 587 Regulates "Speech About Speech"

AB 587 triggers strict scrutiny for the additional reason that it infringes not only X Corp.'s constitutionally protected right to decide what content is permitted on X, but also the public's right to access and disseminate constitutionally protected content on the platform that the State deems objectionable.

Such regulations—those that affect "speech about speech"—have historically triggered heightened First Amendment protections because they create a "triple whammy" of First Amendment problems by impacting not only the speaker's First Amendment rights to make editorial judgments about what speech to permit on a given platform (e.g., a bookstore, school library, movie theater, or social media platform), but also the public's right to access (and in this context, also create) constitutionally protected speech that the government may not suppress directly.

For instance, in *Smith* v. *California*, the Supreme Court struck down, on First Amendment grounds, a Los Angeles ordinance imposing strict liability on booksellers selling obscene books because such liability would "tend to restrict the

public's access to forms of the printed word which the State could not constitutionally suppress directly." 361 U.S. 147, 154 (1958). The Court made clear that it was striking down the law to avoid "self-censorship, compelled by the State, *[which] would be a censorship affecting the whole public*, hardly less virulent for being privately administered." *Id*.

The same concerns that animated the Supreme Court's decision in *Smith* are present here. Like the ordinance in *Smith*, AB 587 is designed to and will likely have the functional effect of pressuring X Corp. and other large social media platforms to change their content-moderation practices. *See, e.g.*, 6-ER-1064; 6-ER-982. This pressure, in turn, will impact what content the hundreds of millions of X users may access and disseminate on the platform.

Given these heightened concerns, courts have not hesitated to apply heightened scrutiny to and strike down laws, like AB 587, that impact not only speakers, but those who act as gatekeepers for the speech of others. In *Software Ass'n* v. *Blagojevich*, for example, the Seventh Circuit applied strict scrutiny to and struck down an Illinois law requiring companies selling video games to identify some games as "sexually explicit" and to distribute brochures explaining the rating system to consumers. 469 F.3d 641, 651–53 (7th Cir. 2006). Like the law in *Blagojevich*, which compelled "subjective" and "opinion-based" disclosures, *id.* at 652, AB 587 forces X Corp. to make potentially billions (*see* 7-ER-1107) of highly

subjective, difficult decisions about controversial content—specifically, whether content on X "belong[s] to any of the categories" set forth in §22677(a)(3).

More recently, the Fifth Circuit followed this rationale in *Wong*, in which it struck down a Texas law requiring schoolbook vendors that do business with public schools to give library material ratings of "sexually explicit," "sexually relevant," or "no rating." 91 F.4th at 325. Like the law in *Wong*, which required the vendors to "undertake contextual analyses, weighing and balancing many factors to determine a rating for each book," *id.* at 340, AB 587 requires X Corp. to make similar judgments with respect to each "item[] of content" on the platform, *see* §22677(a)(5)(A). And, for similar reasons, courts have routinely struck down efforts by states to give legal effect to MPAA ratings to films. *See, e.g.*, *Motion Picture Ass'n of America* v. *Specter*, 315 F.Supp. 824 (E.D. Pa. 1970) (striking down Pennsylvania law prohibiting showing previews for "X" and "R" rated movies at "G" or "GP" films).

By substantively impacting X Corp.'s constitutionally protected speech (its moderation of content), AB 587, like the regulations in *Smith*, *Blagojevic*, *Wong*, and *Specter*, impacts multiple layers of constitutionally protected speech. When regulations such as AB 587 directly or indirectly impact the speech of those charged with making decisions about the proverbial "gates of speech"—e.g., books (*Smith* and *Wong*), video games (*Blagojevic*), and motion pictures (*Specter*)—they

necessarily interfere not only with the speech rights of the gatekeepers, but also those of speakers and listeners on the other side of the gate, thus warranting heightened scrutiny.

C. AB 587 Fails To Survive Any Level Of Scrutiny

1. Strict And Intermediate (*Central Hudson*) Scrutiny

AB 587's TOS Report requirement fails constitutional muster under both strict scrutiny and intermediate scrutiny under *Central Hudson*. It fails strict scrutiny because it is not "narrowly tailored to serve compelling state interests," *Reed*, 576 U.S. at 171, and "less restrictive alternative[s] would serve the Government's purpose," *IMDb.com*, 962 F.3d at 1125.

The TOS Report similarly fails intermediate scrutiny under *Central Hudson* because it does not "directly and materially advance[] a substantial government interest" and is "more extensive than [] necessary to further that interest." *Junior Sports Mags. Inc.* v. *Bonta*, 80 F. 4th 1109, 1116 (9th Cir. 2023). Moreover, under *Central Hudson*, it was the State's "burden to demonstrate that the harms it recites are real and that its restriction will in fact alleviate them to a material degree." *Ibanez* v. *Fla. Dept. of Bus. and Prof'l Regul., Bd. of Acct.*, 512 U.S. 136, 146 (1994). AG Bonta failed to meet this burden. *See, e.g.*, *Italian Colors Rest.* v. *Becerra*, 878 F.3d 1165, 1177 (9th Cir. 2018) (California AG failed to satisfy burden where he "relie[d] solely on the legislative history of [the law] to argue that the California Legislature

understood the [public interest] to be real and adopted [the law] to eliminate that danger.").

First, the governmental interest purportedly served by the TOS Report is neither compelling nor substantial. The State asserts that AB 587 "requir[es] social media companies to be transparent about their content-moderation policies and decisions so that consumers can make informed decisions about where they consume and disseminate news and information." 4-ER-354. However, at least as to X Corp., the TOS Report does not remedy any harm that is "real" rather than "purely hypothetical." *NIFLA*, 585 U.S. at 776.

There is no evidence, for example, that consumers lack or desire more information about how content is regulated on X. Nor could there be, since X provides detailed information to the public about what categories of content are not permitted on X and how such content is regulated. 7-ER-1114–16. Put another way, the State does not have a compelling interest in requiring X Corp. to be transparent about its content-moderation policies, because X Corp. is already transparent about those policies. X Corp. has dedicated immense time, energy, and financial and employee resources to ensuring that its content-moderation policies—including the eight illustrative examples included in the factual record—are accessible and understandable to its consumers. 4-ER-533–94; 7-ER-1115–17.

Moreover, mandatory disclosure is not a compelling governmental interest in the "speech about speech" context presented by AB 587, precisely because laws governing compelled "speech about speech" implicate First Amendment concerns that are simply not present with mandatory transparency laws outside of the speech context. While transparency laws outside of the speech context—e.g., a law requiring restaurants to disclose their Health Department inspection grades—may accomplish their goal of having businesses take certain desired non-speech conduct (e.g., clean up their kitchens) without having any adverse impact on the public's or the restaurants' First Amendment rights, the same cannot be said of a law, like AB 587, that is designed to pressure social media companies to change their content-moderation policies. Indeed, even the authors of AB 587 concede that the law is designed to do precisely that. *See, e.g.*, 5-ER-738; 5-ER-704–05.

Second, even if the State's purported interest were compelling or substantial (and it is not), the TOS Report is not narrowly tailored to further that interest and is also more extensive than necessary to further it. There is no evidence that the law will "directly and materially advance" the goal of allowing consumers to make informed choices about platform use. The TOS Report will likely either (1) do no more than incentivize covered companies to use different content-moderation categories than those in the statute to avoid making the required disclosures or (2)

not compel enough information to enable consumers to make informed decisions about what platforms to use. Either way, the TOS Report may not stand.

The TOS Report is also not narrowly tailored and is more extensive than necessary to further the State's purported goal, because there are several less-restrictive alternatives that the government could have used to promote informed decision-making about social media use. Most obviously, (1) the State could have conducted its own review and published its own findings about the content-moderation policies and practices of social media companies, which are largely already publicly available. *See, e.g.*, 4-ER-533–94. Furthermore, California could have enacted versions of the TOS Report that (2) apply only to companies that do not already disclose their moderation policies; or (3) while requiring covered companies to disclose their content-moderation policies, do not require them to take positions on specific categories of controversial speech.

Accordingly, the TOS Report fails to satisfy strict scrutiny and intermediate scrutiny under *Central Hudson*.

2. *Zauderer*

Finally, the district court erred by finding that the TOS Report would withstand scrutiny under *Zauderer*. According to the district court, the TOS Report's compelled disclosures are not "unduly burdensome" because they do not "effectively rule[] out the speech [they] accompan[y]." 1-ER-6. That holding is

based on a cramped and erroneous interpretation of what constitutes undue burden under *Zauderer*.

Here, although the TOS Report's compelled disclosures may not physically "crowd out" the protected speech of X Corp. and its users as on a product label, they are nonetheless "unduly burdensome" and "[u]njustified" under *Zauderer* because they "chill[] protected speech." *Milavetz*, 559 U.S. at 250. Specifically, the TOS Report—and the threat of enforcement, investigation, and corresponding monetary and injunctive penalty—pressures X Corp. to moderate content in a particular way and thereby interferes with its constitutionally protected editorial judgment. *See NetChoice (Fla.)*, 34 F.4th at 1230–31 (finding SB 7072's "explanation" requirement "unconstitutional under *Zauderer*" because it was "'unduly burdensome' and would 'chill protected speech'—platforms' exercise of editorial judgment").

The TOS Report's compelled disclosures are also unduly burdensome because they chill the protected speech of X platform *users*—both as creators and viewers of speech. *See Volokh*, 656 F.Supp.3d at 445 ("The law requires that social media networks develop policies and procedures with respect to hate speech . . . This could have a profound chilling effect on social media users and their protected freedom of expression.").

Finally, the district court also erred by concluding that AG Bonta carried his burden of showing that the TOS Report's compelled disclosures are "reasonably related to a substantial government interest." 1-ER-6. As set forth above in Section I(A)(iii), the TOS Report does not further any substantial government interest.

II. SECTION 230 PREEMPTS AB 587

The district court also erroneously concluded that AB 587 is not preempted by Section 230 because any penalization of X Corp. for moderating content covered by §230(c)(2) without making the TOS Report's compelled disclosures to the State's satisfaction would not be "liability stemming from a company's content-moderation activities *per se*." 1-ER-8. On that basis, the district court concluded that AB 587 "does not interfere with companies' ability to self-regulate offensive third party content without fear of liability." 1-ER-8.

The district court erred. Given (1) AB 587's vague structure, (2) its hefty penalty provisions, and (3) the fact that AG Bonta has already used threats of enforcement of AB 587 to pressure X Corp. to limit or disfavor content covered by AB 587, *see* 6-ER-1062–63, it is clear that AB 587 facilitates governmental pressure to conform content-moderation decisions to the State's liking, subject to threats of enforcement. Congress enacted Section 230 to broadly protect interactive computer

service providers like X Corp.[8] from exactly this sort of interference and pressure. The district court also failed to acknowledge that Section 230 broadly protects such providers from liability for "*[a]ny action* voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected," 47 U.S.C. §230(c)(2)(A), which necessarily includes such actions taken without the TOS Report's mandated disclosures. For these reasons and those set forth below, AB 587 is preempted by Section 230.

Congress has the power under the Supremacy Clause of the U.S. Constitution to preempt state law when there exists a "conflict" between federal and state law—that is, "where it is impossible to comply with both state and federal requirements, or where state law stands as an obstacle to the accomplishment and execution of the full purpose and objectives of Congress." *Indus. Truck Ass'n, Inc.* v. *Henry*, 125 F.3d 1305, 1309 (9th Cir. 1997). A state law is also preempted by federal law on the basis of "express preemption" when "Congress explicitly defines the extent to which its enactments preempt state law." *Id.* Whether a law is preempted is "almost

[8] X Corp. is an interactive computer service provider under the statute. *See* 47 U.S.C. §230(f)(2).

entirely a question of Congressional intent." *Radici* v. *Associated Ins. Co.*, 217 F.3d 737, 741 (9th Cir. 2000).

47 U.S.C. §230(e)(3), the statute's "express preemption provision[]," *ACA Connects* v. *Bonta*, 24 F.4th 1233, 1246 (9th Cir. 2022), states that "[n]o cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section." Section 230 thus "explicitly preempts inconsistent state laws." *HomeAway.com, Inc.* v. *City of Santa Monica*, 918 F.3d 676, 681 (9th Cir. 2019).

First, AB 587 is expressly preempted by Section 230 because its enforcement provisions allow AG Bonta to bring or threaten to bring a civil action (potentially already after issuing document demands and other requests) against a social media company if he believes it has "omit[ted] or misrepresent[ed]" any reported information. Take a social media company policy prohibiting COVID misinformation on its platform: if the AG believes that reports about COVID originating in a Wuhan lab constitute misinformation (as many did), but the company declines to censor such reports, the AG could bring an investigation or action claiming that the platform "misrepresent[ed]" that it prohibits COVID misinformation or "omit[ted]" an aspect of its misinformation policy. In that way, AB 587 effectively allows the AG to end-run Section 230 and second-guess platforms' decisions under the guise of policing purported misrepresentations or

omissions. That AG Bonta can impose substantial costs on social media companies by either bringing an action or serving investigative discovery demands contravenes the immunity afforded by Section 230, regardless of whether civil penalties are imposed. "[S]ection 230 must be interpreted to protect websites ***not merely from ultimate liability, but from having to fight costly and protracted legal battles***." *Fair Hous. Council of San Fernando Valley* v. *Roommates.com, LLC*, 521 F.3d 1157, 1175 (9th Cir. 2008) (en banc); *id.* at 1174 ("close cases, we believe, must be resolved in favor of immunity").

Moreover, applying §230(c)(2)'s broad immunity here comports with "Congressional intent." *Radici*, 217 F.3d at 741. Congress enacted Section 230 "to encourage ***voluntary monitoring*** for offensive or obscene material." *Carafano* v. *Metrosplash.com, Inc.*, 339 F.3d 1119, 1122 (9th Cir. 2003); *see, e.g.*, *Doe* v. *Internet Brands, Inc.*, 824 F.3d 846, 851–52 (9th Cir. 2016) ("[A] website should be able to act as a 'Good Samaritan' to ***self-regulate*** offensive third-party content without fear of liability."). AB 587 was enacted for the purpose of "pressur[ing]" X Corp. to "eliminate hate speech and disinformation" on its platform (*e.g.*, 5-ER-738), and AG Bonta has already used enforcement of AB 587 to pressure X to change its content-moderation policies to suit the State (*e.g.*, 6-ER-1064). This is entirely antithetical to Section 230's objective to encourage self-regulation unfettered by governmental threats of liability of the sort AB 587 permits and encourages.

Second, AB 587 is also conflict preempted by Section 230. If X Corp. takes actions in good faith to moderate objectionable content, but those actions are deemed to misrepresent or be omitted from X Corp.'s TOS Report, or if the requisite disclosures are deemed otherwise inadequate, X Corp. will face liability. But Section 230(c)(2) protects X Corp. from liability for any such action because such liability falls within the broad scope of immunity intended by Congress, as evidenced by the use of the term "any action." *See, e.g.*, *PC Drivers Headquarters, LP* v. *Malwarebytes Inc.*, 371 F.Supp.3d 652, 660 (N.D. Cal. 2019) (the "any action" language in Section 230 evidences Congressional intent to apply "broad" immunity).

"Any action" means *any* action. It does not permit a state to dictate a required process for good faith moderation actions and impose penalties for noncompliance, especially when the process has both the purpose and likely effect of "pressuring" social media companies to change their content-moderation practices to conform to the State's liking. If there were a law that, for instance, imposed liability on social media content-moderation decisions unless they were made by a panel consisting entirely of registered Republicans, that law would violate Section 230, even though it does not mandate any particular content-moderation outcomes. AB 587 similarly limits social media companies' discretion to moderate content and empowers AG Bonta to second-guess moderation outcomes by claiming that they are inconsistent with the disclosures or insufficiently disclosed. Given the broad scope of Section

230, the State cannot accomplish indirectly (through mandated processes designed to favor certain content-moderation outcomes) what it is forbidden to do directly (impose liability directly for such outcomes).

III. X CORP. SATISFIES THE REMAINING *WINTER* FACTORS

Absent preliminary injunctive relief, X Corp. will continue to suffer irreparable harm, as the "loss of First Amendment freedoms, for even minimal periods of time, unquestionably constitutes irreparable injury." *Roman Cath. Diocese of Brooklyn* v. *Cuomo*, 592 U.S. 14, 19 (2020); *Cal. Chamber of Com.* v. *Council for Educ. and Rsch. on Toxics*, 29 F.4th 468, 482 (9th Cir. 2022). Because X Corp. has established not only a colorable First Amendment claim, but a "likelihood that [AB 587] violates the U.S. Constitution," it has "also established that both the public interest and the balance of the equities favor a preliminary injunction." *Ariz. Dream Act Coal.* v. *Brewer*, 757 F.3d 1053, 1069 (9th Cir. 2014).

CONCLUSION

This Court should reverse the district court's decision denying a preliminary injunction as to AB 587's TOS Report requirement and penalty provisions.

DATED: February 14, 2024

/s/ Joel Kurtzberg

CAHILL GORDON & REINDEL LLP
Joel Kurtzberg
Floyd Abrams
Jason Rozbruch
Lisa J. Cole
32 Old Slip
New York, NY 10005
Telephone: 212-701-3120
Facsimile: 212-269-5420

DOWNEY BRAND LLP
William R. Warne
Meghan M. Baker
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Telephone: 916-444-1000
Facsimile: 916-444-2100

Attorneys for Appellant X Corp.

STATEMENT OF RELATED CASES

Pursuant to Ninth Circuit Rule 28-2.6, X Corp. is unaware of any related cases currently pending in this Court.

DATED: February 14, 2024

/s/ Joel Kurtzberg

CAHILL GORDON & REINDEL LLP
Joel Kurtzberg
Floyd Abrams
Jason Rozbruch
Lisa J. Cole
32 Old Slip
New York, NY 10005
Telephone: 212-701-3120
Facsimile: 212-269-5420

DOWNEY BRAND LLP
William R. Warne
Meghan M. Baker
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Telephone: 916-444-1000
Facsimile: 916-444-2100

Attorneys for Appellant X Corp.

CERTIFICATE OF COMPLIANCE FOR BRIEFS

This brief complies with the word length limits permitted by Ninth Circuit Rule 32-1, as it contains 14,000 words, excluding the items exempted by Federal Rule of Appellate Procedure 32(f).

This brief complies with the type size and typeface requirements under Federal Rules of Appellate Procedure 32(a)(5) and (6), because it has been prepared in a proportionately spaced typeface using Microsoft Word Times New Roman 14-point font.

DATED: February 14, 2024

/s/ Joel Kurtzberg

CAHILL GORDON & REINDEL LLP
Joel Kurtzberg
Floyd Abrams
Jason Rozbruch
Lisa J. Cole
32 Old Slip
New York, NY 10005
Telephone: 212-701-3120
Facsimile: 212-269-5420

DOWNEY BRAND LLP
William R. Warne
Meghan M. Baker
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Telephone: 916-444-1000
Facsimile: 916-444-2100

Attorneys for Appellant X Corp.

CERTIFICATE OF SERVICE

I hereby certify that I caused this document to be electronically filed on this date with the Clerk of the Court for the United States Court of Appeals for the Ninth Circuit using the Appellate Electronic Filing system. All registered case participants will be served via the Appellate Electronic Filing system.

DATED: February 14, 2024

/s/ Joel Kurtzberg

CAHILL GORDON & REINDEL LLP
Joel Kurtzberg
Floyd Abrams
Jason Rozbruch
Lisa J. Cole
32 Old Slip
New York, NY 10005
Telephone: 212-701-3120
Facsimile: 212-269-5420

DOWNEY BRAND LLP
William R. Warne
Meghan M. Baker
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Telephone: 916-444-1000
Facsimile: 916-444-2100

Attorneys for Appellant X Corp.

ADDENDUM

Assembly Bill No. 587

CHAPTER 269

An act to add Chapter 22.8 (commencing with Section 22675) to Division 8 of the Business and Professions Code, relating to social media.

[Approved by Governor September 13, 2022. Filed with Secretary of State September 13, 2022.]

LEGISLATIVE COUNSEL'S DIGEST

AB 587, Gabriel. Social media companies: terms of service.

Existing law requires an operator of a commercial website or online service that collects personally identifiable information through the internet about individual consumers residing in California who use or visit its commercial website or online service to make its privacy policy available to consumers, as specified.

This bill would require a social media company, as defined, to post their terms of service for each social media platform, as defined, owned or operated by the company in a specified manner and with additional specified information, subject to certain exceptions. The bill would define "terms of service" to mean a policy or set of policies adopted by a social media company that specifies, at least, the user behavior and activities that are permitted on the internet-based service owned or operated by the social media company, and the user behavior and activities that may subject the user or an item of content to being actioned, as defined.

This bill would also require the social media company to submit reports, as specified, starting no later than January 1, 2024, to the Attorney General. The bill would specify the information required by the reports, including, but not limited to, the current version of the terms of service for each social media platform owned or operated by the company, specified categories of content and what policies the social media company has for that platform to address that content, and data related to violations of the terms of service for each platform. The bill would require the Attorney General to make all terms of service reports submitted pursuant to those provisions available to the public in a searchable repository on its official internet website.

The bill would state the intent of the Legislature that a social media company that violates the above provisions shall be subject to meaningful remedies sufficient to induce compliance with these provisions, and would specify civil penalties that a company shall be liable for if the bill's provisions are violated, and how the Attorney General or a city attorney may bring an action against violators. The bill would specify that the duties, obligations, remedies, and penalties imposed by the bill are cumulative to existing law.

The people of the State of California do enact as follows:

SECTION 1. It is the intent of the Legislature that a social media company that violates this act shall be subject to meaningful remedies sufficient to induce compliance with this act.

SEC. 2. Chapter 22.8 (commencing with Section 22675) is added to Division 8 of the Business and Professions Code, to read:

Chapter 22.8. Content Moderation Requirements for Internet Terms of Service

22675. For purposes of this chapter, the following definitions apply:

(a) "Actioned" means a social media company, due to a suspected or confirmed violation of the terms of service, has taken some form of action, including, but not limited to, removal, demonetization, deprioritization, or banning, against the relevant user or relevant item of content.

(b) (1) "Content" means statements or comments made by users and media that are created, posted, shared, or otherwise interacted with by users on an internet-based service or application.

(2) "Content" does not include media put on a service or application exclusively for the purpose of cloud storage, transmitting files, or file collaboration.

(c) "Public or semipublic internet-based service or application" excludes a service or application used to facilitate communication within a business or enterprise among employees or affiliates of the business or enterprise, provided that access to the service or application is restricted to employees or affiliates of the business or enterprise using the service or application.

(d) "Social media company" means a person or entity that owns or operates one or more social media platforms.

(e) "Social media platform" means a public or semipublic internet-based service or application that has users in California and that meets both of the following criteria:

(1) (A) A substantial function of the service or application is to connect users in order to allow users to interact socially with each other within the service or application.

(B) A service or application that provides email or direct messaging services shall not be considered to meet this criterion on the basis of that function alone.

(2) The service or application allows users to do all of the following:

(A) Construct a public or semipublic profile for purposes of signing into and using the service or application.

(B) Populate a list of other users with whom an individual shares a social connection within the system.

(C) Create or post content viewable by other users, including, but not limited to, on message boards, in chat rooms, or through a landing page or main feed that presents the user with content generated by other users.

(f) "Terms of service" means a policy or set of policies adopted by a social media company that specifies, at least, the user behavior and activities that are permitted on the internet-based service owned or operated by the social media company, and the user behavior and activities that may subject the user or an item of content to being actioned.

22676. (a) A social media company shall post terms of service for each social media platform owned or operated by the company in a manner reasonably designed to inform all users of the social media platform of the existence and contents of the terms of service.

(b) The terms of service posted pursuant to subdivision (a) shall include all of the following:

(1) Contact information for the purpose of allowing users to ask the social media company questions about the terms of service.

(2) A description of the process that users must follow to flag content, groups, or other users that they believe violate the terms of service, and the social media company's commitments on response and resolution time.

(3) A list of potential actions the social media company may take against an item of content or a user, including, but not limited to, removal, demonetization, deprioritization, or banning.

(c) The terms of service posted pursuant to subdivision (a) shall be available in all Medi-Cal threshold languages, as defined in subdivision (c) of Section 128552 of the Health and Safety Code, in which the social media platform offers product features, including, but not limited to, menus and prompts.

22677. (a) On a semiannual basis in accordance with subdivision (b), a social media company shall submit to the Attorney General a terms of service report. The terms of service report shall include, for each social media platform owned or operated by the company, all of the following:

(1) The current version of the terms of service of the social media platform.

(2) If a social media company has filed its first report, a complete and detailed description of any changes to the terms of service since the previous report.

(3) A statement of whether the current version of the terms of service defines each of the following categories of content, and, if so, the definitions of those categories, including any subcategories:

(A) Hate speech or racism.

(B) Extremism or radicalization.

(C) Disinformation or misinformation.

(D) Harassment.

(E) Foreign political interference.

(4) A detailed description of content moderation practices used by the social media company for that platform, including, but not limited to, all of the following:

(A) Any existing policies intended to address the categories of content described in paragraph (3).

(B) How automated content moderation systems enforce terms of service of the social media platform and when these systems involve human review.

(C) How the social media company responds to user reports of violations of the terms of service.

(D) How the social media company would remove individual pieces of content, users, or groups that violate the terms of service, or take broader action against individual users or against groups of users that violate the terms of service.

(E) The languages in which the social media platform does not make terms of service available, but does offer product features, including, but not limited to, menus and prompts.

(5) (A) Information on content that was flagged by the social media company as content belonging to any of the categories described in paragraph (3), including all of the following:

(i) The total number of flagged items of content.

(ii) The total number of actioned items of content.

(iii) The total number of actioned items of content that resulted in action taken by the social media company against the user or group of users responsible for the content.

(iv) The total number of actioned items of content that were removed, demonetized, or deprioritized by the social media company.

(v) The number of times actioned items of content were viewed by users.

(vi) The number of times actioned items of content were shared, and the number of users that viewed the content before it was actioned.

(vii) The number of times users appealed social media company actions taken on that platform and the number of reversals of social media company actions on appeal disaggregated by each type of action.

(B) All information required by subparagraph (A) shall be disaggregated into the following categories:

(i) The category of content, including any relevant categories described in paragraph (3).

(ii) The type of content, including, but not limited to, posts, comments, messages, profiles of users, or groups of users.

(iii) The type of media of the content, including, but not limited to, text, images, and videos.

(iv) How the content was flagged, including, but not limited to, flagged by company employees or contractors, flagged by artificial intelligence software, flagged by community moderators, flagged by civil society partners, and flagged by users.

(v) How the content was actioned, including, but not limited to, actioned by company employees or contractors, actioned by artificial intelligence software, actioned by community moderators, actioned by civil society partners, and actioned by users.

(b) (1) A social media company shall electronically submit a semiannual terms of service report pursuant to subdivision (a), covering activity within the third and fourth quarters of the preceding calendar year, to the Attorney General no later than April 1 of each year, and shall electronically submit

a semiannual terms of service report pursuant to subdivision (a), covering activity within the first and second quarters of the current calendar year, to the Attorney General no later than October 1 of each year.

(2) Notwithstanding paragraph (1), a social media company shall electronically submit its first terms of service report pursuant to subdivision (a), covering activity within the third quarter of 2023, to the Attorney General no later than January 1, 2024, and shall electronically submit its second terms of service report pursuant to subdivision (a), covering activity within the fourth quarter of 2023, to the Attorney General no later than April 1, 2024. A social media platform shall submit its third report no later than October 1, 2024, in accordance with paragraph (1).

(c) The Attorney General shall make all terms of service reports submitted pursuant to this section available to the public in a searchable repository on its official internet website.

22678. (a) (1) A social media company that violates the provisions of this chapter shall be liable for a civil penalty not to exceed fifteen thousand dollars ($15,000) per violation per day, and may be enjoined in any court of competent jurisdiction.

(2) A social media company shall be considered in violation of the provisions of this chapter for each day the social media company does any of the following:

(A) Fails to post terms of service in accordance with Section 22676.

(B) Fails to timely submit to the Attorney General a report required pursuant to Section 22677.

(C) Materially omits or misrepresents required information in a report submitted pursuant to Section 22677.

(3) In assessing the amount of a civil penalty pursuant to paragraph (1), the court shall consider whether the social media company has made a reasonable, good faith attempt to comply with the provisions of this chapter.

(b) Actions for relief pursuant to this chapter shall be prosecuted exclusively in a court of competent jurisdiction by the Attorney General or by a city attorney of a city having a population in excess of 750,000, or by a city attorney in a city and county in the name of the people of the State of California upon their own complaint or upon the complaint of a board, officer, person, corporation, or association.

(c) If an action pursuant to this section is brought by the Attorney General, one-half of the penalty collected shall be paid to the treasurer of the county in which the judgment was entered, and one-half to the General Fund. If the action is brought by a city attorney, one-half of the penalty collected shall be paid to the treasurer of the city in which the judgment was entered, and one-half to the treasurer of the county in which the judgment was entered.

22679. (a) The duties and obligations imposed by this chapter are cumulative to any other duties or obligations imposed under local, state, or federal law and shall not be construed to relieve any party from any duties or obligations imposed under law.

(b) The remedies or penalties provided by this chapter are cumulative to each other and to any other remedies or penalties available under local, state, or federal law.

22680. This chapter shall not apply to a social media company that generated less than one hundred million dollars ($100,000,000) in gross revenue during the preceding calendar year.

22681. This chapter shall not be construed to apply to an internet-based service or application for which interactions between users are limited to direct messages, commercial transactions, consumer reviews of products, sellers, services, events, or places, or any combination thereof.

O

Editorial Notes

REFERENCES IN TEXT

The Communications Assistance for Law Enforcement Act, referred to in subsecs. (a) and (e), is title I of Pub. L. 103–414, Oct. 25, 1994, 108 Stat. 4279, which is classified generally to subchapter I (§ 1001 et seq.) of chapter 9 of this title. For complete classification of this Act to the Code, see Short Title note set out under section 1001 of this title and Tables.

This chapter, referred to in subsecs. (d) and (e)(2), was in the original "this Act", meaning act June 19, 1934, ch. 652, 48 Stat. 1064, known as the Communications Act of 1934, which is classified principally to this chapter. For complete classification of this Act to the Code, see section 609 of this title and Tables.

§ 230. Protection for private blocking and screening of offensive material

(a) Findings

The Congress finds the following:

(1) The rapidly developing array of Internet and other interactive computer services available to individual Americans represent an extraordinary advance in the availability of educational and informational resources to our citizens.

(2) These services offer users a great degree of control over the information that they receive, as well as the potential for even greater control in the future as technology develops.

(3) The Internet and other interactive computer services offer a forum for a true diversity of political discourse, unique opportunities for cultural development, and myriad avenues for intellectual activity.

(4) The Internet and other interactive computer services have flourished, to the benefit of all Americans, with a minimum of government regulation.

(5) Increasingly Americans are relying on interactive media for a variety of political, educational, cultural, and entertainment services.

(b) Policy

It is the policy of the United States—

(1) to promote the continued development of the Internet and other interactive computer services and other interactive media;

(2) to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by Federal or State regulation;

(3) to encourage the development of technologies which maximize user control over what information is received by individuals, families, and schools who use the Internet and other interactive computer services;

(4) to remove disincentives for the development and utilization of blocking and filtering technologies that empower parents to restrict their children's access to objectionable or inappropriate online material; and

(5) to ensure vigorous enforcement of Federal criminal laws to deter and punish trafficking in obscenity, stalking, and harassment by means of computer.

(c) Protection for "Good Samaritan" blocking and screening of offensive material

(1) Treatment of publisher or speaker

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.

(2) Civil liability

No provider or user of an interactive computer service shall be held liable on account of—

(A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or

(B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).[1]

(d) Obligations of interactive computer service

A provider of interactive computer service shall, at the time of entering an agreement with a customer for the provision of interactive computer service and in a manner deemed appropriate by the provider, notify such customer that parental control protections (such as computer hardware, software, or filtering services) are commercially available that may assist the customer in limiting access to material that is harmful to minors. Such notice shall identify, or provide the customer with access to information identifying, current providers of such protections.

(e) Effect on other laws

(1) No effect on criminal law

Nothing in this section shall be construed to impair the enforcement of section 223 or 231 of this title, chapter 71 (relating to obscenity) or 110 (relating to sexual exploitation of children) of title 18, or any other Federal criminal statute.

(2) No effect on intellectual property law

Nothing in this section shall be construed to limit or expand any law pertaining to intellectual property.

(3) State law

Nothing in this section shall be construed to prevent any State from enforcing any State law that is consistent with this section. No cause of action may be brought and no liability may be imposed under any State or local law that is inconsistent with this section.

(4) No effect on communications privacy law

Nothing in this section shall be construed to limit the application of the Electronic Communications Privacy Act of 1986 or any of the amendments made by such Act, or any similar State law.

(5) No effect on sex trafficking law

Nothing in this section (other than subsection (c)(2)(A)) shall be construed to impair or limit—

(A) any claim in a civil action brought under section 1595 of title 18, if the conduct underlying the claim constitutes a violation of section 1591 of that title;

---

[1] So in original. Probably should be "subparagraph (A)."

(B) any charge in a criminal prosecution brought under State law if the conduct underlying the charge would constitute a violation of section 1591 of title 18; or

(C) any charge in a criminal prosecution brought under State law if the conduct underlying the charge would constitute a violation of section 2421A of title 18, and promotion or facilitation of prostitution is illegal in the jurisdiction where the defendant's promotion or facilitation of prostitution was targeted.

(f) Definitions

As used in this section:

(1) Internet

The term ''Internet'' means the international computer network of both Federal and non-Federal interoperable packet switched data networks.

(2) Interactive computer service

The term ''interactive computer service'' means any information service, system, or access software provider that provides or enables computer access by multiple users to a computer server, including specifically a service or system that provides access to the Internet and such systems operated or services offered by libraries or educational institutions.

(3) Information content provider

The term ''information content provider'' means any person or entity that is responsible, in whole or in part, for the creation or development of information provided through the Internet or any other interactive computer service.

(4) Access software provider

The term ''access software provider'' means a provider of software (including client or server software), or enabling tools that do any one or more of the following:

(A) filter, screen, allow, or disallow content;

(B) pick, choose, analyze, or digest content; or

(C) transmit, receive, display, forward, cache, search, subset, organize, reorganize, or translate content.

(June 19, 1934, ch. 652, title II, § 230, as added Pub. L. 104–104, title V, § 509, Feb. 8, 1996, 110 Stat. 137; amended Pub. L. 105–277, div. C, title XIV, § 1404(a), Oct. 21, 1998, 112 Stat. 2681–739; Pub. L. 115–164, § 4(a), Apr. 11, 2018, 132 Stat. 1254.)

Editorial Notes

REFERENCES IN TEXT

The Electronic Communications Privacy Act of 1986, referred to in subsec. (e)(4), is Pub. L. 99–508, Oct. 21, 1986, 100 Stat. 1848, as amended. For complete classification of this Act to the Code, see Short Title of 1986 Amendment note set out under section 2510 of Title 18, Crimes and Criminal Procedure, and Tables.

CODIFICATION

Section 509 of Pub. L. 104–104, which directed amendment of title II of the Communications Act of 1934 (47 U.S.C. 201 et seq.) by adding section 230 at end, was executed by adding the section at end of part I of title II of the Act to reflect the probable intent of Congress and amendments by sections 101(a), (b), and 151(a) of Pub. L. 104–104 designating §§ 201 to 229 as part I and adding parts II (§ 251 et seq.) and III (§ 271 et seq.) to title II of the Act.

AMENDMENTS

2018—Subsec. (e)(5). Pub. L. 115–164 added par. (5).

1998—Subsec. (d). Pub. L. 105–277, § 1404(a)(3), added subsec. (d). Former subsec. (d) redesignated (e).

Subsec. (d)(1). Pub. L. 105–277, § 1404(a)(1), inserted ''or 231'' after ''section 223''.

Subsecs. (e), (f). Pub. L. 105–277, § 1404(a)(2), redesignated subsecs. (d) and (e) as (e) and (f), respectively.

Statutory Notes and Related Subsidiaries

EFFECTIVE DATE OF 2018 AMENDMENT

Pub. L. 115–164, § 4(b), Apr. 11, 2018, 132 Stat. 1254, provided that: ''The amendments made by this section [amending this section] shall take effect on the date of the enactment of this Act [Apr. 11, 2018], and the amendment made by subsection (a) shall apply regardless of whether the conduct alleged occurred, or is alleged to have occurred, before, on, or after such date of enactment.''

EFFECTIVE DATE OF 1998 AMENDMENT

Amendment by Pub. L. 105–277 effective 30 days after Oct. 21, 1998, see section 1406 of Pub. L. 105–277, set out as a note under section 223 of this title.

SAVINGS

Pub. L. 115–164, § 7, Apr. 11, 2018, 132 Stat. 1255, provided that: ''Nothing in this Act [see Short Title of 2018 Amendment note set out under section 1 of Title 18, Crimes and Criminal Procedure] or the amendments made by this Act shall be construed to limit or preempt any civil action or criminal prosecution under Federal law or State law (including State statutory law and State common law) filed before or after the day before the date of enactment of this Act [Apr. 11, 2018] that was not limited or preempted by section 230 of the Communications Act of 1934 (47 U.S.C. 230), as such section was in effect on the day before the date of enactment of this Act.''

SENSE OF CONGRESS

Pub. L. 115–164, § 2, Apr. 11, 2018, 132 Stat. 1253, provided that: ''It is the sense of Congress that—

''(1) section 230 of the Communications Act of 1934 (47 U.S.C. 230; commonly known as the 'Communications Decency Act of 1996') was never intended to provide legal protection to websites that unlawfully promote and facilitate prostitution and websites that facilitate traffickers in advertising the sale of unlawful sex acts with sex trafficking victims;

''(2) websites that promote and facilitate prostitution have been reckless in allowing the sale of sex trafficking victims and have done nothing to prevent the trafficking of children and victims of force, fraud, and coercion; and

''(3) clarification of such section is warranted to ensure that such section does not provide such protection to such websites.''

Executive Documents

EXECUTIVE ORDER NO. 13925

Ex. Ord. No. 13925, May 28, 2020, 85 F.R. 34079, which related to moderation of content posted on social media platforms, was revoked by Ex. Ord. No. 14029, § 1, May 14, 2021, 86 F.R. 27025.

§ 231. Restriction of access by minors to materials commercially distributed by means of World Wide Web that are harmful to minors

(a) Requirement to restrict access

(1) Prohibited conduct

Whoever knowingly and with knowledge of the character of the material, in interstate or foreign commerce by means of the World Wide Web, makes any communication for commercial purposes that is available to any minor and that includes any material that is harmful to minors shall be fined not more than $50,000, imprisoned not more than 6 months, or both.

(2) Intentional violations

In addition to the penalties under paragraph (1), whoever intentionally violates such paragraph shall be subject to a fine of not more than $50,000 for each violation. For purposes of this paragraph, each day of violation shall constitute a separate violation.

(3) Civil penalty

In addition to the penalties under paragraphs (1) and (2), whoever violates paragraph (1) shall be subject to a civil penalty of not more than $50,000 for each violation. For purposes of this paragraph, each day of violation shall constitute a separate violation.

(b) Inapplicability of carriers and other service providers

For purposes of subsection (a), a person shall not be considered to make any communication for commercial purposes to the extent that such person is—

(1) a telecommunications carrier engaged in the provision of a telecommunications service;

(2) a person engaged in the business of providing an Internet access service;

(3) a person engaged in the business of providing an Internet information location tool; or

(4) similarly engaged in the transmission, storage, retrieval, hosting, formatting, or translation (or any combination thereof) of a communication made by another person, without selection or alteration of the content of the communication, except that such person's deletion of a particular communication or material made by another person in a manner consistent with subsection (c) or section 230 of this title shall not constitute such selection or alteration of the content of the communication.

(c) Affirmative defense

(1) Defense

It is an affirmative defense to prosecution under this section that the defendant, in good faith, has restricted access by minors to material that is harmful to minors—

(A) by requiring use of a credit card, debit account, adult access code, or adult personal identification number;

(B) by accepting a digital certificate that verifies age; or

(C) by any other reasonable measures that are feasible under available technology.

(2) Protection for use of defenses

No cause of action may be brought in any court or administrative agency against any person on account of any activity that is not in violation of any law punishable by criminal or civil penalty, and that the person has taken in good faith to implement a defense authorized under this subsection or otherwise to restrict or prevent the transmission of, or access to, a communication specified in this section.

(d) Privacy protection requirements

(1) Disclosure of information limited

A person making a communication described in subsection (a)—

(A) shall not disclose any information collected for the purposes of restricting access to such communications to individuals 17 years of age or older without the prior written or electronic consent of—

(i) the individual concerned, if the individual is an adult; or

(ii) the individual's parent or guardian, if the individual is under 17 years of age; and

(B) shall take such actions as are necessary to prevent unauthorized access to such information by a person other than the person making such communication and the recipient of such communication.

(2) Exceptions

A person making a communication described in subsection (a) may disclose such information if the disclosure is—

(A) necessary to make the communication or conduct a legitimate business activity related to making the communication; or

(B) made pursuant to a court order authorizing such disclosure.

(e) Definitions

For purposes of this subsection,[1] the following definitions shall apply:

(1) By means of the World Wide Web

The term ''by means of the World Wide Web'' means by placement of material in a computer server-based file archive so that it is publicly accessible, over the Internet, using hypertext transfer protocol or any successor protocol.

(2) Commercial purposes; engaged in the business

(A) Commercial purposes

A person shall be considered to make a communication for commercial purposes only if such person is engaged in the business of making such communications.

(B) Engaged in the business

The term ''engaged in the business'' means that the person who makes a communication, or offers to make a communication, by means of the World Wide Web, that includes any material that is harmful to minors, devotes time, attention, or labor to such activities, as a regular course of such person's trade or business, with the objective of earning a profit as a result of such activities (although it is not necessary that the person make a profit or that the making or offering to make such communications be the per-

[1] So in original. Probably should be ''section,''.