UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF CALIFORNIA
SACRAMENTO DIVISION

CHRISTOPHER KOHLS, ET AL.,

Plaintiffs,

v.

ROBERT A. BONTA, ET AL.,

Defendants.

Case No. 2:24-cv-02527-JAM-CKD

PLAINTIFFS' RESPONSE TO DEFENDANTS' MOTION FOR SUMMARY JUDGMENT ON AB 2655

EXTENDED ORAL ARGUMENT REQUESTED

Date: August 5, 2025
Time: 1:00 p.m.
Crtrm: 6, 14th floor
Judge: John A. Mendez

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases **Page(s)**

Statutes

INTRODUCTION

As Defendants concede, AB 2655 "regulates on the basis of content and is subject to strict scrutiny"—a "demanding" standard. Defs.' Br. (Doc. 49-2) at 16–17. But Defendants downplay the burden imposed on them by strict scrutiny. For AB 2655 to survive, Defendants must ***prove*** that it furthers the government's interest in "protecting free and fair elections," is narrowly tailored to achieve that interest, and that ***no less speech-restrictive alternatives*** would do so. "[T]he danger of censorship presented by a facially content-based statute requires that that weapon be employed only where it is ***necessary*** to serve the asserted [compelling] interest." *R.A.V.* v. *City of St. Paul*, 505 U.S. 377, 395 (1992) (emphasis in original).[1] Defendants fall far short of making that showing.

Attempting their heavy lift, Defendants offer the Declaration of R. Michael Alvarez, Ph.D, which discusses the supposed "growing problem" of political content generated using artificial intelligence ("AI"). But as stated in the Expert Declaration of Christopher Lucas, Ph.D., the limited scholarship in this new field offers no evidence that political deepfakes are unique, compared to other types of false political speech. That means the State lacks a valid basis for singling out political deepfakes for special treatment (as AB 2655 does), compared to other types of false political speech. In addition, the State has not met its burden of demonstrating that there are no less-speech-restrictive alternatives to AB 2655's censorship regime. As explained in the Expert Declaration of John W. Ayers, Ph.D., counterspeech methods, such as X's Community Notes (*see, e.g.*, Pls.' Br. (Doc. 46-1) at 20–23), not only work well, they better accomplish the goal of preserving free and fair elections than do top-down censorship regimes, like AB 2655.

AB 2655 fails strict scrutiny because there is no proof that the law will be effective in ensuring that elections are free and fair. The real purpose of the statute seems to be to impermissibly replace the content-moderation policies of the covered platforms with one of the government's own choosing. But there is no evidence that the existing policies of the covered platforms are not working. While California speculates that deepfakes have the ***potential*** to deceive people and impact elections, there is no evidence whatsoever that even a single "deceptive deepfake" that

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

would be covered by the statute caused actual harm to anyone or undermined the integrity of an election in any way. Such speculation cannot satisfy strict scrutiny.

Nor is there proof that AB 2655 would do a good job of removing the truly deceptive content without over-censoring truthful and protected content. Defendants argue that the law is narrowly tailored because it does not apply to certain types of protected speech, such as satire and parody, *e.g.*, Defs.' Br. at 13, but this misses the point entirely. The law sets up a system of enforcement that incentivizes over-censorship by the platforms because it (1) imposes substantial costs for censoring or labeling too little, (2) imposes no costs whatsoever for censoring or labeling too much, and (3) uses politically biased candidates and elected officials as enforcers of the law. The effects of this enforcement regime are obvious, do not rely on any disputed issues of material fact, and cannot be reconciled with the First Amendment. The Court should thus grant summary judgment to Plaintiffs and deny summary judgment to Defendants. The undisputed record evidence makes clear that the statutory scheme will result in the chilling and censorship of substantial swaths of constitutionally protected, political speech.

AB 2655 also violates and is preempted by the immunity afforded to interactive computer service providers under 47 U.S.C. §§ 230(c)(1) and 230(c)(2). Contrary to Defendants' position, overwhelming authority holds that Section 230 covers suits for purely injunctive relief, such as those enabled by AB 2655. Defendants also advance the position that, if a platform's editorial discretion is protected by the First Amendment, the platform cannot also be protected under Section 230. But, with good reason, such an interpretation has never been accepted within this Circuit. It would nullify Section 230's plain language and intent, contradicting Ninth Circuit case law.

Finally, AB 2655 is overbroad and void for vagueness. Top-to-bottom, the statute's requirements are expansive and inherently subjective, forcing platforms, in condensed windows of time, to determine whether content is "false[]," "satire or parody," and "reasonably likely" to "harm the reputation or electoral prospects of a candidate" or "falsely undermine confidence in the outcome of one or more election contests." Against the backdrop of a regime that immunizes entirely taking down or labeling content, imposes severe costs for ***not*** taking down or labeling content, and authorizes politically motivated individuals to sue platforms with unfettered discretion,

it is clear that AB 2655 cannot withstand constitutional muster.

ARGUMENT

I. AB 2655 Violates Free Speech Rights.

a. AB 2655 is Facially Invalid.

Plaintiffs appropriately bring a facial challenge. AB 2655, in all of its applications, attempts to regulate speech. Defendants concede this. *See* Defs.' Br. at 16 ("Defendants do not dispute that AB 2655 implicates the expressive activity of online platforms such as plaintiffs . . . under *Moody*, 603 U.S. at 727, platforms have an expressive interest in the expressive 'selection, ordering, and labeling of third party posts.'"). And they concede that AB 2655 regulates that speech based on content. *Id.*

California relies primarily on *Moody* v. *NetChoice, LLC* to suggest that facial challenges always require an analysis of overbreadth. 603 U.S. 707, 723–24 (2024); Defs.' Br. at 12. But *Moody* explained that overbreadth is merely a tool to scrutinize a law's scope. *See* 603 U.S. at 740 ("In the usual First Amendment case, [courts] must decide whether to apply strict or intermediate scrutiny."). When a facial challenge attacks a law that "raise[s] the same First Amendment issues" for all regulated speakers, courts should proceed to apply the relevant level of constitutional scrutiny. *X Corp.* v. *Bonta*, 116 F.4th 888, 898–99 (9th Cir. 2024). Because AB 2655 facially discriminates against speech based on content and viewpoint, at a minimum, it is subject to strict scrutiny. *Id.* at 899–903 (sustaining facial challenge against California law that interfered with social media platforms' editorial discretion); *Reed* v. *Town of Gilbert, Ariz.*, 576 U.S. 155, 165 (2015) ("A law that is content based on its face is subject to strict scrutiny regardless . . . "). It also fails *Moody*'s test for facial overbreadth with a sweeping effect on wide swaths of protected, core political speech.

First, because the law prohibits content likely to "harm" candidates and election officials, but not content likely to help, no matter how false that content might be (*see, e.g.*, Pls.' Br. at 4, 24, 28–29), AB 2655 discriminates based on viewpoint, which, under controlling precedent, makes it automatically facially invalid. *Iancu* v. *Brunetti*, 588 U.S. 388, 398 (2019) (striking down

prohibition of "immoral or scandalous" trademarks as viewpoint-discriminatory); *Matal* v. *Tam*, 582 U.S. 218, 243 (2017) (striking down policy denying trademarks that disparage or bring into contempt any living or dead people or groups as viewpoint-discriminatory); *R.A.V.,* 505 U.S. at 380, 393–94 (striking down criminal ordinance prohibiting the display of symbols that "arouse[] anger, alarm, or resentment in others on the basis of race, color, creed, religion, or gender" as viewpoint-discriminatory). AB 2655 targets speech ***critical*** of political candidates—e.g., in the words of the statute, "that is reasonably likely to harm the reputation or electoral prospects of a candidate," Section 20513(a)(2)(A)—but not speech that ***helps*** a candidate's reputation or electoral prospects undeservedly. For example, if a candidate posted AI-generated videos falsely depicting her as buying homes for homeless people or rescuing stray animals in her district, those videos would not be covered by AB 2655, despite the fact that they have the potential to mislead voters and undermine the integrity of elections. Only content likely to harm candidates is covered by the statute. This is classic viewpoint discrimination. *Iancu*, *Matal*, and *R.A.V.*, all of which involved facially unconstitutional laws that targeted critical or "negative" speech but provided dissimilar treatment for speech that was "positive," are indistinguishable from AB 2655.

Defendants ignore AB 2655's viewpoint discrimination, and that omission is fatal. As the Supreme Court has made clear, a law that is viewpoint-discriminatory cannot be "salvage[d] by virtue of its constitutionally permissible applications." *Iancu*, 588 U.S. at 398. Put differently, a "finding of viewpoint bias end[s] the matter," *id*. at 399, because a viewpoint-discriminatory law has no valid applications. *Iancu* confirmed that the Court has "never applied that kind of analysis"—i.e., the test that Defendants focus on exclusively: assessing overbreadth by examining whether the law's "unconstitutional applications are [] substantial relative to the statute's plainly legitimate sweep"—to "a viewpoint-discriminatory law." *Id.* at 398–99; *see also Matal*, 582 U.S. at 248 (Kennedy, J., concurring) (a finding of viewpoint discrimination "renders unnecessary any extended treatment of other questions raised by the parties"); *R.A.V.*, 505 U.S. at 381 (striking

down, as facially unconstitutional, a viewpoint-discriminatory ordinance without engaging in an overbreadth analysis).

This case is thus materially distinct from *Moody* vis-à-vis the appropriateness of a facial challenge, given that the Texas and Florida laws in *Moody*, unlike AB 2655, were not viewpoint-discriminatory speech restrictions on their face. *See NetChoice, LLC* v. *Att'y Gen., Fla.*, 34 F.4th 1196, 1224 (11th Cir. 2022) ("[W]e don't think that . . . the entire Act is impermissibly viewpoint-based" because "when a statute is facially constitutional, a plaintiff cannot bring a free-speech challenge by claiming that the lawmakers who passed it acted with a constitutionally impermissible purpose"). This case is also distinct from *Moody*, where there were significant questions about the scope of the Texas and Florida laws, including whether they applied to services like "direct messaging," "events management," "online marketplace[s]," "financial exchanges" or even "ride-sharing," some of which may not have implicated the platforms' free expression rights. *See Moody*, 603 U.S. at 724–25. There are no such questions about the scope of AB 2655's applications, given that the law only applies to speech on "[l]arge online platform[s]," as defined by Section 20512(h). AB 2655 thus applies in the same manner to every such large online platform, "rais[ing] the same First Amendment issues" for each of them. *X Corp.*, 116 F.4th at 899.

Second, while the *Moody* overbreadth standard doesn't apply because AB 2655 targets speech based on content and viewpoint, the law would still flunk *Moody*'s test for facial validity. Defendants attempt to offer constitutional applications of AB 2655, but these efforts in fact show that the law's unconstitutional applications substantially outweigh its constitutional ones, rendering it overbroad. Defendants assert that "digitally created photos of then-former President Donald Trump being arrested"—like the one Plaintiff Rickert wants to post—constitute "examples of potential political deepfakes." DSUF (Doc. 49-5) ¶ 19; PSUF (Doc. 47) ¶¶ 212–13. They also discuss "images that allegedly showed then-former President Trump embracing Dr. Anthony Fauci." Defs' Br. at 3. These images make political statements and merit the utmost First Amendment protection. Rickert understood the messages to communicate that, "This is the strategy

to get Kamala Harris elected—political retaliation." PSUF ¶ 90. Yet, as Defendants acknowledge, AB 2655 would require large online platforms to remove them.

Defendants also emphasize that AB 2655 does not apply to outlandish parody or satire, arguing that this fact defeats overbreadth. *See* Defs.' Br. at 13 (offering examples like swapping "President Trump's face onto the body of Superman in a movie clip"). But these hypothetical examples of supposed parody or satire ring hollow. The evidence in the record makes clear that, even if the statute exempts far-fetched parody and satire, it will still have a serious chilling effect on core political speech.

Take the example of Plaintiff Kohls's "Kamala Harris Ad PARODY" video, *see* Pls.' Br. at 18, which is obviously a parody and even includes the word "PARODY" in its title. Governor Newsom publicly stated that this ad parody "should be illegal," and, when AB 2655 was passed, he stated that he "just signed a bill to make this illegal in the state of California." *See* PSUF ¶ 151 (Kurtzberg Decl. Ex. 13). The point is not that Plaintiffs are right about the video being parody and that the Governor is wrong; rather, the point is that, even if the video qualifies as parody under the law, AB 2655 will deter content creators from posting protected parody and satire and incentivize platforms to remove that content because politicians can easily ***claim*** that the video is "illegal" under the statute and impose considerable costs merely by suing to have the content labeled or taken down. *See Susan B. Anthony List* v. *Driehaus*, 573 U.S. 149, 163 (2014) (it "misses the point" to argue that mens rea limitation protects speakers when enforcement depends on others' interpretations). To show how the law would apply to Kohls' video, Defendants now emphasize that the video was reposted without the parody disclaimer. DSUF ¶ 3. In other words, AB 2655 chills protected speech by requiring content creators and platforms to either censor themselves or defend close calls in court at their own great expense. The result is censorship of important and protected political speech, even if it arguably does not fall within the statute. Moreover, that concern is particularly acute given that AB 2655 regulates online platforms, because platforms—which host hundreds of millions or billions of posts daily—often have little context when assessing

whether a particular post is parody or satire. *See* PSUF ¶¶ 184–85. Platforms will therefore be highly incentivized to err on the side of censoring content at scale to avoid liability risk.

Defendants also argue that the definition of "materially deceptive content" under the statute "perfectly tracks the line drawn under the First Amendment for when satire or parody can be actionable as defamation." Defs.' Br. at 13. Not so. AB 2655 doesn't just apply to speech "***that damages*** a candidate's electoral chances or reputation." Defs.' Br. at 14. It applies to speech that "***that is reasonably likely to harm***" it. § 20513(a)(2)(A).

The Court emphasized this delta in preliminarily enjoining AB 2839 (a substantially similar law to AB 2655 aimed at content creators, rather than platforms) as likely violating the First Amendment on its face. *See Kohls* v. *Bonta*, 752 F. Supp. 3d 1187, 1193 (E.D. Cal. 2024). AB 2655, like AB 2839, "does much more than punish potential defamatory statements since the statute does not require actual harm and sanctions any digitally manipulated content that is 'reasonably likely' to 'harm' the amorphous 'electoral prospects' of a candidate or elected official[.]" *Id.* Defendants' argument, that AB 2655 targets "false statement[s] made in association with legally cognizable harm or for the purpose of material gain," Defs.' Br. at 14–15 (citing *Animal Legal Def. Fund* v. *Wasden*, 878 F.3d 1184, 1199 (9th Cir. 2018)), thus fails as well.

Third, Defendants vastly underestimate how AB 2655 targets protected speech. They argue that, "[e]ven if an occasional parody or satire might meet the definition of 'materially deceptive content,' it is still not subject to AB 2655 because the statute expressly exempts satire or parody." Defs.' Br. at 16. But AB 2655 chills speech, even when platforms succeed in defending close calls, because the entire enforcement system imposes costs on them only for resisting censorship and labeling. In addition, Defendants impermissibly attempt to reverse-engineer a category of unprotected speech by simply exempting one category of protected speech. Pls.' Br. at 26–27. The same goes for Defendants' argument that a "platform is not subject to suit when it lacks the requisite awareness that a piece of media has been digitally manipulated." *Id.* at 27. Again, even if a court ultimately determines that a platform is not liable under the statute because it did not have sufficient awareness that the content was materially deceptive, platforms are nonetheless "subject to suit"

under Sections 20515(b) and 20516, along with all of the burden and costs that attend even unmeritorious litigation, which raise First Amendment problems on their own. "[T]he real potential damage is done at the time a complaint is filed." *281 Care Comm.* v. *Arneson*, 766 F.3d 774, 792 (8th Cir. 2014) (striking down, on First Amendment grounds, a state statute criminalizing certain false speech in connection with ballot questions); *see also Susan B. Anthony List*, 573 U.S. at 163.

In the end, Defendants fail to come to grips with the profound chilling effect that AB 2655 will have on wide swaths of constitutionally protected, core political speech. AB 2655 creates a slanted censorship regime that incentivizes platforms to remove or label any materials presenting even a close call under the statute. That alone will result in the censorship and/or labeling of substantial amounts of protected and important political speech, not to mention that it impermissibly substitutes the government's judgment about what is "materially deceptive" speech for the constitutionally protected judgments of the covered platforms. Such censorship regimes violate the First Amendment. *See NetChoice, LLC* v. *Bonta*, 113 F.4th 1101, 1118 (9th Cir. 2024) (striking down law on First Amendment grounds that "deputizes covered businesses into serving as censors for the State"); *Smith* v. *Cal.*, 361 U.S. 147, 154 (1959) (government efforts to impose a system that encourages "self-censorship" by booksellers violates the First Amendment because it results in censorship "compelled by the State . . . affecting the whole public"); *Nat'l Rifle Ass'n* v. *Vullo*, 602 U.S. 175, 190 (2024) ("a government official cannot do indirectly what she is barred from doing directly"). This result will occur uniformly to all covered platforms in all applications of the law, making AB 2655 facially unconstitutional.

b. AB 2655 is Invalid as-Applied to Plaintiffs.

Defendants concede, as did two California legislature Committees (Pls.' Br. at 11–12), that AB 2655 "regulates on the basis of content and is subject to strict scrutiny." Defs.' Br. at 16.

AB 2655 also triggers strict scrutiny for the independent reason that the Labeling and Reporting Requirements compel speech. Pls.' Br. at 28. Defendants argue that the Labeling Requirement should be subject only to "exacting scrutiny" because the Ninth Circuit has recognized "a lower standard for disclosures in political speech." Defs.' Br. at 24 (citing *Smith* v. *Helzer*, 95

F.4th 1207, 1214 (9th Cir. 2024)). That is wrong. Transparency laws that compel speech still trigger strict scrutiny. *Nat'l Inst. of Fam. & Life Advocs.* v. *Becerra ("NIFLA")*, 585 U.S. 755, 766 (2018) (licensing notices); *Riley* v. *Nat'l Fed'n of the Blind of N.C., Inc.*, 487 U.S. 781, 797–98 (1988) ("compelled statements of 'fact'" for fundraisers); *X Corp.*, 116 F.4th at 902 ("Even a pure 'transparency' measure, if it compels non-commercial speech, is subject to strict scrutiny" (citing *Riley*, 487 U.S. 796–97)). California's line of "exacting scrutiny" cases does not apply for at least three reasons. *First,* campaign finance cases involve "campaign-related expenditures and contributions." *Cal. Pro-Life Council, Inc.* v. *Getman*, 328 F.3d 1088, 1104 (9th Cir. 2003). They allow "the public to 'follow the money.'" *Hum. Life of Wash. Inc.* v. *Brumsickle*, 624 F.3d 990, 997 (9th Cir. 2010). *Second,* AB 2655's compelled disclaimer does not inform voters about "the person or group who is speaking." *Citizens United* v. *FEC*, 558 U.S. 310, 368 (2010). Instead, it "compel[s] speakers to utter" a "particular message," making it "subject to the same rigorous scrutiny" as other content-based laws. *Turner Broad. Sys., Inc.* v. *FCC*, 512 U.S. 622, 642 (1994). *Third*, a "key premise" of campaign-finance laws is that they "may burden the ability to speak, [but] they . . . do not prevent anyone from speaking." *Washington Post* v. *McManus*, 944 F.3d 506, 516 (4th Cir. 2019). Yet the practical effect of AB 2655's Labeling Requirement is to chill much speech in conjunction with the Blocking Requirement. Even so, the Labeling Requirement fails "exacting scrutiny" for the reasons discussed below.

Defendants cannot meet their strict-scrutiny burden because they have not shown and cannot show that the statute (1) is "narrowly tailored to serve compelling state interests," *NIFLA*, 585 U.S. at 766, and (2) that no "less restrictive alternative would serve the [g]overnment's purpose," *X Corp.*, 116 F.4th at 903; *see also* Pls.' Br. at 30–33.

i. *The State Cannot Meet its Burden of Showing that AB 2655 is Narrowly Tailored to Further a Compelling Government Interest.*

a. AB 2655 fails to advance a compelling state interest.

Defendants argue that AB 2655 serves a compelling government interest in "protecting free and fair elections." Defs.' Br. at 17. But choosing such a lofty interest comes at a steep cost—that is, the difficulty of demonstrating that the statute is the least speech-restrictive means of furthering

such an interest. *See, e.g.*, *McIntyre* v. *Ohio Elections Comm'n*, 514 U.S. 334, 347–48, 357 (1995) (striking down statute prohibiting anonymous pamphleteering for lack of tailoring); *281 Care Comm.*, 766 F.3d at 787–96 (striking down Minnesota political misinformation statute for want of tailoring). Here, AB 2655 does not advance California's stated interest because (1) the law doesn't target a "real" harm; (2) it codifies a highly paternalistic approach to democracy; and (3) California has no evidence that "materially deceptive content" causes its perceived harms.

First, California cannot show that AB 2655 addresses a problem that is "real," rather than "purely hypothetical." *NIFLA*, 585 U.S. at 776. The State points to the California Legislature's findings that (1) "fake images or audio or video content" can "skew election results" and "undermine trust in the ballot counting process"; (2) "with current available technology, '[v]oters will not know what images, audio, or video they can trust'"; and (3) "[s]uch manipulated content can 'prevent voters from voting and deceive voters based on fraudulent content.'" Defs.' Br. at 17. Dr. Alvarez opines that "research and studies confirm" these "legislative findings"—i.e., that "political deepfakes have proliferated online and can influence voters' behavior, choices, and trust in the electoral process and electoral outcomes." *Id.* at 18.

However, the most current academic literature in this area—cited with approval by Defendants' own expert (Alvarez Decl. ¶ 15 (citing research of Plaintiffs' expert, Christopher Lucas))—demonstrates that there is nothing remarkable about political deepfakes, as compared to other types of false political speech. Pls.' Response to Defendants' Statement of Undisputed Facts ("PRDSUF") ¶ 23 (Lucas Decl. ¶¶ 13–15). Specifically, concerns about political deepfakes and their potential to introduce new waves of misinformation are speculative, as there is little evidence that deepfake videos are notably more persuasive than conventional forms of misinformation (e.g., written text). *Id.* ¶ 23 (Lucas Decl. ¶ 14). For instance, while Dr. Alvarez asserts that deepfake videos can be convincing and "sticky" and can spread quickly (Alvarez Decl. ¶¶ 15, 22, 25), the evidence he cites in support of these propositions consists primarily of studies about misinformation in written text. PRDSUF ¶ 23 (Lucas Decl. ¶ 13); *see also id.* (Ayers Decl. ¶ 47) (stating that he could "not identify a single study that examined the duration of political deepfake effects relative

to other forms of misinformation"); *id*. ¶¶ 18–19 (Ayers Decl. ¶ 26) ("nearly all of Alvarez's claims about deepfakes are speculative. . . . [H]is claims are based less on deepfake research and more on speculation surrounded by citations to unrelated research."). Dr. Alvarez thus fails to provide material evidence supporting the premise that deepfakes are considerably more damaging than existing forms of misinformation. *See also* Alvarez Decl. ¶ 11 ("political deepfakes ***might*** affect a viewer"); *id.* ¶ 16 ("the effects of political deepfakes on voter trust and confidence in elections are understudied"). There is also no evidence in the record of any actual harm caused by a single "deepfake" that would be covered by AB 2655. There is therefore no valid basis for singling out political deepfakes for special treatment (as compared to other types of false political speech), as AB 2655 does, or for concluding that, even if deepfakes are a problem, that AB 2655 will be effective in addressing that problem.

Defendants also attempt to justify AB 2655 by pointing to the Political Deepfakes Incidents Database ("PDID"), which Defendants claim likely understates the number of political deepfakes online. *See* Defs.' Br. at 3–4; Defs.' Statement of Undisputed Facts ("DSUF") ¶ 17. But as Plaintiffs' expert Dr. Ayers makes clear, the PDID does not support Defendants' arguments—it undermines them. The very database showcased by Defendants as illustrating the existence of the supposed problem of political deepfakes demonstrates that almost all of the political deepfakes in the database are either not covered by AB 2655 or were identified as being AI-generated through various counterspeech methods that are being consistently deployed, effectively, to address viewer deception. Ayers Decl. ¶¶ 53–82; *id*. at Ex. 2 (spreadsheet showing analysis). No record evidence demonstrates that a single "deepfake" in the PDID undermined the integrity of an election or influenced a single voter in any way.

The PDID actually demonstrates that counterspeech has been extremely effective in identifying and uncovering deepfakes. As of April 24, 2025, with respect to 99.9% of all engagements with non-duplicate posts in the PDID and 94.1% of all non-duplicate posts in the PDID: (1) AB 2655 did not apply to the content in the post to begin with, e.g., because the content was not materially deceptive or did not portray a political candidate covered by the statute, and/or

(2) an effective form of counterspeech had been employed—i.e., the content was already disclosed as a deepfake, AI, or as otherwise manipulated, by (a) the poster of the content, (b) Community Notes, (c) Grok AI, (d) fact checkers employed by some covered platforms, and/or (d) user comments.[2] PRDSUF ¶ 17 (Ayers Decl. ¶¶ 55, 69, 72, 79, 80); *see also* Lucas Decl. ¶ 9 ("there is no evidence that banning political deepfakes will mitigate th[e] issue" of individuals "dismiss[ing] genuine evidence that challenges their pre-existing beliefs"); *Kohls*, 752 F. Supp. 3d at 1191 ("Especially as to political speech, counter speech is the tried and true buffer and elixir, not speech restriction."). The PDID also includes substantial amounts of content that even Defendants would concede would ***not*** be covered by AB 2655, further undermining any argument that the PDID supports the statute's constitutionality. *Compare* Defs.' Br. at 13 (Trump as Superman image example), *with* PRDSUF ¶ 18 (Ayers Decl. ¶ 63) (showing Trump as Captain America image in the PDID). Dr. Alvarez himself has even publicly advocated for the use of automated tools that he claims can be employed at scale to identify misinformation online. *Id.* ¶ 24 (Ayers Decl. ¶ 50).

Defendants' repeated citation of the well-covered, AI-generated, fake robocalls before the 2024 New Hampshire primary, telling voters not to go to the polls, *see, e.g.*, Defs.' Br. at 17–18, does not demonstrate that the problem that AB 2655 is designed to address is "real." For one thing, there is no evidence in the record that any covered platform played any role in the alleged robocalls, so AB 2655 would not even apply to such conduct. Indeed, and perhaps unsurprisingly, the PDID contains nothing similar to the New Hampshire robocall example. And, in any event, as the evidence in the record demonstrates, the creators of any such content would likely be punished through already-existing speech-neutral criminal laws of the type that have been used to punish the perpetrators of the supposed Biden robocall scheme. *See* Sec. I(b)(ii) below.

Second, courts reject California's "highly paternalistic approach [of] limiting what people may hear." *Eu* v. *San Francisco Cnty. Democratic Cent. Comm.*, 489 U.S. 214, 223 (1989) (affirming *San Francisco Cnty. Democratic Cent. Comm.* v. *Eu*, 826 F.2d 814 (9th Cir. 1987)).

[2] The comparable figure for engagements with non-duplicative posts on the X platform in the PDID is 99.9%, and the figure for non-duplicative X posts in the PDID is 97.5%. Ayers Decl. ¶¶ 79–80.

The First Amendment demands "greater faith in the ability of individual voters to inform themselves" and to decide who ought to be the party nominee. 826 F.2d at 836.

Here, too, California's "claim that it is enhancing the ability of its citizenry to make wise decisions by restricting the flow of information to them must be viewed with some skepticism." *Id.* at 836. California could have written a statute targeting legally cognizable harms, like "fraud and corruption." *Eu*, 489 U.S. at 229. Instead, it targeted certain messages, worrying they might "deceive[]" voters, "prevent" citizens from voting, or undermine their trust in elections. Defs.' Br. at 17. But free and open debate does not "pose a risk to democracy." *Contra id.* at 18. It sustains democracy. Poking fun at politicians and "critici[zing] [the] government is at the very center of the constitutionally protected area of free discussion." *Chaker* v. *Crogan*, 428 F.3d 1215, 1227 (9th Cir. 2005). "California's ban on" certain political speech "is a form of paternalism that is inconsistent with the First Amendment." *Eu*, 826 F.2d at 836.

Third, California has failed to show that the proscribed digitally created or modified content causes any harms justifying burdens on First Amendment rights. California's similar feature in *Brown* v. *Ent. Merchs. Ass'n*, 564 U.S. 786 (2011), is instructive. To justify its ban on sales of violent video games to minors, California cited studies showing violent games were "significantly linked to increases in aggressive behaviour[.]" *Video Software Dealers Ass'n* v. *Schwarzenegger*, 556 F.3d 950, 963 (9th Cir. 2009). But evidence of "correlation" wasn't enough to satisfy the First Amendment because it did not "prove that violent video games cause minors to act aggressively." *Brown*, 564 U.S. at 800.

There's no proof of causation here either. California frets that the most nefarious content (and that is only part of what AB 2655 covers) "*can* influence voters," "*could* undermine the public's trust in elections," or "*may* alter voters' behavior." Defs.' Br. at 17–18. Alvarez too recognizes that problems from "political deepfakes" "*may be* profound," they "*might* affect a viewer," "*can* introduce uncertainty," "*can* sow confusion," "*can* change subsequent behavior," etc. Alvarez Decl. ¶¶ 9, 11, 12. He acknowledges that the research is nascent, and "the effects of political deepfakes on voter trust and confidence in elections are understudied." *Id.* ¶¶ 14, 16

(explaining engagement with conspiracy theories "is associated with lower turnout in elections"). In short, California offers only "mere speculation about serious harms." *Progressive Democrats for Soc. Just.* v. *Bonta*, 73 F.4th 1118, 1126 (9th Cir. 2023). That is not enough. *Id.*

On this record, Defendants have failed to satisfy their burden to demonstrate that the problem they are attempting to address with AB 2655 is real and "not purely hypothetical." *NIFLA*, 585 U.S. at 776 (government bears burden to make this showing).

b. Defendants cannot show AB 2655 is narrowly tailored.

Defendants' narrow tailoring arguments fare no better. They look to several features: (1) AB 2655 applies only to "materially deceptive content" (Defs.' Br. at 19–20); (2) it imposes a mens rea requirement (*id*. at 20); (3) the Blocking and Labeling Requirements require no affirmative action by covered platforms (*id*. at 20–21); (4) the Blocking and Labeling Requirements are subject to temporal windows that supposedly reflect the risk of harm that each is designed to address (*id.* at 21–22); and (5) what candidates personally say or do purportedly will have the largest impact on voter behavior (*id.*). None of these features saves the statute or narrowly tailors it.

First, Defendants argue that AB 2655 narrowly regulates only "materially deceptive content," exempting parody or satire. But, as discussed above, AB 2655 still requires platforms to censor or label swaths of protected speech, like the images of then-candidate Trump being arrested. *See* Sec. I(a) above. Defendants' arguments also ignore that the statute—because it entirely immunizes over-censorship and over-labeling, while potentially imposing real costs on platforms that do not censor or label enough within a highly compressed timeframe (*see, e.g.*, Pls.' Br. at 2, 12, 14–15, 42)—will, in practice, result in the censorship of much content at the margins that is not "materially deceptive content," including content that is satire or parody. *See, e.g.*, *id.* at 19 (Khara Boender of the CCIA explaining that AB 2655 will resemble the takedown regime under the Digital Millennium Copyright Act, where platforms are "required to block content almost constantly in order to ensure compliance"); Sec. III below (discussing Governor Newsom's statements indicating that content that is parody or satire is nonetheless prohibited by AB 2655). Thus, even if AB 2655

covers only "materially deceptive content," that does not demonstrate narrow tailoring, because in practice the statute will significantly affect content that falls outside that category.

Second, Defendants assert that the statute is tailored through its mens rea requirement—i.e., that a platform need only act as to content as to which it "knows or acts with reckless disregard for the fact that [it] meets the requirements of [each] section." Defs.' Br. at 20. This argument suffers from a similar flaw. Given the volume of content reported to platforms, the statute's compressed timeframes (during which it will be "next-to-impossible" for platforms to determine whether content is covered by the statute, *see* Pls.' Br. at 41), and, again, that the statute entirely immunizes over-censorship and over-labeling. It does nothing "to realistically stop the potential for abuse of [AB 2655's] mechanism." *281 Care Comm.*, 766 F.3d at 794. It "does not eliminate the threat posed by [the statute's] staggering reach." *Animal Legal Def. Fund*, 878 F.3d at 1197–98.

Third, Defendants argue that the statute is narrowly tailored, because "AB 2655 does not require an online platform to affirmatively look for and proactively remove or label *any* content." Defs.' Br. at 20 (emphasis in original). That, however, is untrue. AB 2655 explicitly requires platforms to affirmatively "identify" and "remove" content that is "identical or substantially similar" to materially deceptive content previously removed under the statute. § 20513(c). The word "identify" clearly imposes an independent, affirmative obligation on the platforms. And that burden is substantial, given how subjective it is to try to determine whether certain content is "substantially similar" to other content.

Fourth, Defendants claim the Blocking and Labeling Requirements "work in tandem as a sliding scale," with varying temporal limitations for each requirement. Defs.' Br. at 21. The applicable timeframes for each, however, are still too long. For instance, Defendants attempt to justify the broad application of the Blocking Requirement—which spans 60 days after the election for content that "depicts or pertains to elections officials" (§ 20513(e)(2))—because it applies during the "period after an election while results are being certified." Defs.' Br. at 22. But Defendants' proffered rationale makes no sense, given that "elections officials" often must

complete the certification of votes much earlier, and no later than 38 days after the election. *See* Cal. Elec. Code §§ 15372, 15503, 15504, 15505.

Fifth, Defendants argue that "what a candidate personally says or does is likely to have the biggest impact on a voter's decision-making process." Defs.' Br. at 21–22. But there is no evidentiary support for this debatable assertion, and it was California's burden to provide it. *See NIFLA*, 585 U.S. at 766. Voting is an individual multivariable decision motivated by a variety of factors—not only the candidate, but also party affiliation, incumbent, party platform, and relevant election issues, to name a few. The State has provided no evidence showing the motivations of voters or what influences those motivations.

Finally, the law also flunks strict scrutiny because it is severely underinclusive. It exempts online platforms with fewer than 1 million users and certain broadcasters (despite there being no principled basis to do so, *see Moody*, 603 U.S. at 719, 738); it only applies to "materially deceptive content" and not to other deceptive political speech (a distinction likewise lacking a principled basis, *see* PRDSUF ¶ 18 (Lucas Decl. ¶ 12)); and it does not apply to false speech that makes it look like a candidate for elective office, elections official, or elected official did something positive that he or she did not do (*see* §§ 20513(a)(2), 201514(a)(2)), even though such speech could undermine the integrity of elections. For example, if a candidate for elective office, an elections official, or an elected official actively participated in election fraud and distributed AI-generated videos falsely showing them having done things to prevent the very fraud they participated in, AB 2655 would not apply, despite the fact that such content would pose at least as much risk to the integrity of elections as any "materially deceptive content" covered by the statute.

ii. *AB 2655 Fails Strict Scrutiny Because Less-Speech-Restrictive Alternatives Would Further the State's Goal.*

Nor is AB 2655 the "least restrictive means available for advancing [the State's] interest." *Kohls*, 752 F. Supp. 3d at 1195 (quoting *NetChoice, LLC*, 113 F.4th at 1121). The "First Amendment does not permit speech-restrictive measures when the state may remedy the problem by implementing or enforcing laws that do not infringe on speech." *Id.* The government has the

burden to consider and prove ineffective any "plausible, less restrictive alternative[s]." *U.S.* v. *Playboy Ent. Grp.*, 529 U.S. 803, 816 (2000); *see also Brewer* v. *City of Albuquerque*, 18 F.4th 1205, 1246 (10th Cir. 2021). Here, less-speech-restrictive alternatives that compel a finding of AB 2655's constitutional infirmity are the various (1) effective counterspeech and other methods of Plaintiffs and other covered platforms; (2) initiatives that California could take that don't require censorship; and (3) legal remedies that already address the supposed harms targeted by the statute.

First, as previewed above in Section I(b)(i) and as stated in Plaintiffs' Opening Brief (at 20–23, 24–28; PSUF ¶¶ 114–23, 125–45, 147), the covered platforms already maintain numerous extensive procedures and policies to protect against user deception from content covered by the statute. These procedures and policies further an open, robust marketplace of ideas and provide transparency about controversial or deceptive speech. And even Dr. Alvarez suggests—in a public interview, but not in his declaration—that certain automated tools can now be effectively employed "at scale" to combat misinformation online. PRDSUF ¶¶ 14, 16, 22, 24–26 (Ayers Decl. ¶ 50).

Because counterspeech works better than censorship, the current systems of Plaintiffs and other covered platforms will better achieve the government's goal in protecting free and fair elections than AB 2655 will. Censorship fails, for instance, because the "public cannot have a conversation about potentially misleading speech if content is deleted." *Id.* ¶¶ 26–28 (Ayers Decl. ¶ 22). Nor can the public learn from counterspeech and dialogue about misinformation if the underlying "content is already censored," which would "undermin[e] evidence-based communication and the development of an informed populace." *Id.* ¶¶ 26–27 (Ayers Decl. ¶ 22); *see also id.* ¶ 26 (Ayers Decl. ¶ 21) ("public conversations are often the best way to discover the truth"). In addition, "censorship strategies," such as the one imposed by AB 2655, "risk amplifying misinformation's impact via deepfakes suppressing content." *Id.* ¶¶ 26–28 (Ayers Decl. ¶ 21); *see also id.* ¶ 27 (Lucas Decl. ¶¶ 20–24) (explaining that a ban on political deepfakes will be ineffective, e.g., because of the difficulties in defining and detecting them). Over-censorship without transparency—AB 2655's product—would deprive the electorate of robust political debate needed for free and fair elections.

Defendants do not even attempt to bear their burden of showing that these procedures and policies are inadequate to address the dangers of political deepfakes. That alone is reason to hold that AB 2655 fails strict scrutiny. In fact, the State's own evidence demonstrates the proliferation and effectiveness of the covered platforms' already-existent counterspeech methods. *See id.* ¶ 17 (Ayers Decl. ¶¶ 55, 69, 72, 79, 80) (explaining that as of April 24, 2025, with respect to 99.9% of all engagements with non-duplicate posts in the PDID, 99.9% of all engagements with non-duplicate X posts in the PDID, 97.5% of all non-duplicate X posts in the PDID, and 94.1% of all non-duplicate posts in the PDID, AB 2655 likely would not apply to the content to begin with, and/or an effective form of counterspeech has been employed as to the content).

Evidence used by the State to argue that the "detrimental impact of political deepfakes" is "not merely hypothetical," Defs.' Br. at 3, further proves this point. For instance, the 2023 "digitally created video of former President Biden allegedly announcing the beginning of World War III and the return of the draft," was addressed via a Community Note highlighting that "[t]his presenter retweeted an AI video without including the written warning that this has been created by AI like the original creator did." *See* Liska Decl. (Doc. 49-4) Ex. 14 at 155 of 204. Same for Defendants' example of "deepfake images of then-former President Donald Trump being arrested." Defs.' Br. at 3. Defendants omit that "[a]nnotations from Community Notes" were "attached to [the] tweets to include the context that the Trump Images were AI-generated." Liska Decl. (Doc. 49-4) Ex. 16 at 170 of 204. No record evidence whatsoever suggests that any of these "deepfakes" actually deceived voters, impacted their decisions, or undermined the integrity of elections.

In contrast with the absence of any record evidence that "materially deceptive content" inflicted harm during the 2024 election cycle, the record shows that X effectively labeled potentially misleading content under its Authenticity Policy. Specifically, from August 5, 2024, to February 3, 2025, X labeled 3,700 posts under its Authenticity Policy and its predecessor version. Pls.' Br. at 21; PSUF ¶ 127. And Rumble removes content from its platform that "[p]romotes, supports, or incites violence or unlawful acts." PSUF ¶ 134(c).

Indeed, nothing is stopping the California government—in particular (1) AG Bonta (@AGRobBonta; joined X June 2018), (2) Secretary Weber (@DrWeber4CA; joined X January 2014), (3) Gavin Newsom (@GavinNewsom; joined X December 2007), or (4) their staffs (using those X accounts or other accounts that meet the contributor requirements, *see* Pls.' Br. at 20–21)—from utilizing Community Notes or from leaving posts or replies on particular pieces of content, to inform X users that particular content is "materially deceptive" and/or generated using artificial intelligence and is potentially misleading. X Corp. S&O Decl. ¶ 24. While Defendants argue that it is "hardly feasible" for the government to identify and debunk deceptive digitally altered content, Defs.' Br. at 23, they have not even tried to do so and have no proof that their counterspeech, along with that of other users, will be inadequate to address the problems targeted by the statute.

Second, California has many less restrictive—and more effective means—to advance its stated interest. California could consider an educational campaign to improve digital and media literacy. *44 Liquormart, Inc.* v. *Rhode Island*, 517 U.S. 484, 507 (1996). That would likely be more effective (and would definitely be more constitutional) than censoring speech. "Research . . . shows that promoting digital and media literacy and improving political knowledge can effectively combat misinformation through deepfakes." Lucas Decl. ¶ 30; *see also* Liska Decl. (Doc. 49-4), Ex. 16 at 164 of 204 (citing news article quoting "experts" opining that "the best way to combat visual misinformation is better public awareness and education"). This includes "teaching the importance of source credibility, editorial standards, and the differences between reputable journalism and sensationalist content." PRDSUF ¶¶ 24–26 (Lucas Decl. ¶ 30). It additionally avoids the "false sense of security" of expecting every fake video to have a proscribed disclaimer attached to it. Lucas Decl. ¶ 23.

Alvarez agrees: "with strengthened media literacy skills . . . people can be more likely to identify political deepfakes and less likely to believe that they are accurate[.]" Alvarez Decl. ¶ 40. His only objection is that this "would require a large investment of resources." *Id.* But "[t]he First Amendment does not permit the State to sacrifice speech for efficiency." *NIFLA*, 585 U.S. at 775 (holding "public-information campaign" was less-restrictive alternative and suggesting "California

spent insufficient resources on the advertising campaign"). California fails to show "that it seriously undertook to address the problem with less intrusive tools readily available to it." *McCullen* v. *Coakley*, 573 U.S. 464, 494 (2014).

Third, as this Court determined in its decision preliminarily enjoining AB 2839, statutory causes of action "such as privacy torts, copyright infringement, or defamation already provide recourse" for the supposed harms that AB 2655 seeks to address. *Kohls*, 752 F. Supp. 3d at 1195–96; *see also IMDb.com Inc.* v. *Becerra*, 962 F.3d 1111, 1126 (9th Cir. 2020) ("Because the State has various other laws at its disposal that would allow it to achieve its stated interests while burdening little or no speech, it fails to show that the law is the least restrictive means to protect its compelling interest. That failure alone dooms [the law]."); *Ex parte Stafford*, 2024 WL 4031614, at *4–6 (Tex. Crim. App. Sept. 4, 2024) (striking down, under strict scrutiny, Texas statute prohibiting "knowingly represent[ing] in a campaign communication that the communication emanates from a source other than its true source" because there were "narrower means of achieving the State interests," including enforcing an existing statute).

Specifically, various causes of action exist under state and federal law that would offer relief for the supposed harms AB 2655 seeks to address—including, but not limited to, those for defamation; statutory and common law right of publicity; false light; violations of the Telephone Consumer Protection Act; violations of the Voting Rights Act; and copyright infringement. In addition, the California Elections Code already has numerous provisions that protect against the same harms purportedly targeted by AB 2655.[3]

[3] *See* the following provisions of the California Elections Code:

- *id.* § 20010(a) ("[A] person, firm, association, corporation, campaign committee, or organization shall not, with actual malice, produce, distribute, publish, or broadcast campaign material that contains (1) a picture or photograph of a person or persons into which the image of a candidate for public office is superimposed or (2) a picture or photograph of a candidate for public office into which the image of another person or persons is superimposed.");
- *id.* § 18543(a) ("Every person who knowingly challenges a person's right to vote without probable cause or on fraudulent or spurious grounds, or who engages in mass, indiscriminate, and groundless challenging of voters solely for the purpose of preventing voters from voting or to delay the process of voting, or who fraudulently advises any person that he or she is not eligible to vote or is not registered to vote when in fact that person is eligible or is registered, or who violates Section 14240, is punishable by imprisonment[.]");
- *id.* § 18502(a) ("Any person who in any manner interferes with the officers holding an election or conducting a canvass, or with the voters lawfully exercising their rights of voting at an election, as to prevent the election or canvass from being fairly held and lawfully conducted, is punishable by imprisonment[.]");
- *id.* § 18302(b) ("A person is guilty of a misdemeanor who, with actual knowledge and intent to deceive, causes to

That the harms supposedly targeted by AB 2655 are better left for review by already-existent laws that do not trifle with the First Amendment is exemplified by the aftermath of the "New Hampshire robocall" example at the center of Defendants' motion papers, as the perpetrators of that scheme have faced extensive penalties at the state and federal levels.[4] While Defendants argue—without a single citation in support—that "existing causes of action, such as defamation law, do not adequately address the harms that AB 2655 seeks to prevent," Defs' Br. at 23, they provide no ***evidence*** to support this *ipse dixit* assertion, as the First Amendment requires. There is,

---

be distributed or distributes, including distribution by mail, radio or television broadcast, telephone call, text message, email, or any other electronic means, including over the Internet, literature or any other form of communication to a voter that includes any of the following: (1) The incorrect location of a vote center, office of an elections official, satellite office of an elections official where voting is permitted, vote by mail ballot drop box, or vote by mail ballot drop-off location. (2) False or misleading information regarding the qualifications to vote or to register to vote. (3) False or misleading information regarding the qualifications to apply for, receive, or return a vote by mail ballot. (4) False or misleading information regarding the date of an election or the days, dates, or times voting may occur at a place described in paragraph (1).");

- *id.* § 20500 ("The provisions of Part 2 (commencing with Section 43) of Division 1 of the Civil Code, relating to libel and slander, are fully applicable to any campaign advertising or communication.");
- *id.* § 20501(a) ("A candidate or state measure proponent is liable for any slander or libel committed by a committee that is controlled by that candidate or state measure proponent as defined by Section 82016 of the Government Code if the candidate or state measure proponent willfully and knowingly directs or permits the libel or slander.");
- *id.* § 10.5 ("[P]rimary missions of the Office of Elections Cybersecurity" are "To coordinate efforts between the Secretary of State and local elections officials to reduce the likelihood and severity of cyber incidents that could interfere with the security or integrity of elections in the state."; and "To monitor and counteract false or misleading information regarding the electoral process that is published online or on other platforms and that may suppress voter participation or cause confusion and disruption of the orderly and secure administration of elections.");
- *id.* § 18320(b)-(c) ("It is unlawful for a person, with intent to mislead, deceive, or defraud, to commit an act of political cyberfraud[.]");
- *id.* § 18350(a)(2) ("A person is guilty of a misdemeanor who, with intent to mislead the voters in connection with his or her campaign" or "in connection with the campaign of another person for nomination or election to a public office . . . [a]ssumes, pretends, or implies, by his or her statements, conduct, or campaign materials, that he or she is or has been acting in the capacity of a public officer when that is not the case.");
- *id.* § 18573 (prohibiting fraud that causes voters to miscast their vote).

[4] Steven Kramer, the political consultant who arranged the robocalls received by New Hampshire residents, was indicted on 13 felony counts of voter suppression and 13 misdemeanor counts of impersonation of a candidate in New Hampshire state court in May 2024. PRDSUF ¶ 19 (Kurtzberg II Decl. Ex. 1); Liska Decl. (Doc. 49-4) Ex. 18. As of June 6, 2025, Kramer's state court criminal trial is still ongoing. PRDSUF ¶ 19 (Kurtzberg II Decl. Ex. 2). Kramer was also fined $6 million in September 2024 by the Federal Communications Commission ("FCC") for his role in the robocalls. *Id.* (Kurtzberg II Decl. Ex. 3). And Lingo Telecom, which transmitted the robocalls and incorrectly labeled them with the highest level of caller ID attestation, agreed to pay $1 million to the FCC in August 2024. *Id.* (Kurtzberg II Decl. Ex. 4). Finally, Kramer, Lingo Telecom, and Life Corporation, the company that broadcasted the robocalls, were sued in the District of New Hampshire for: (1) intimidating, threatening, or coercing voters in violation of the Voting Rights Act; (2) disseminating unlawful artificial or prerecorded-voice telephone calls in violation of the Telephone Consumer Protection Act; (3) delivering prerecorded political messages without required disclosure in violation of New Hampshire law; and (4) knowingly misrepresenting the origin of telephone calls containing information about candidates or parties in violation of New Hampshire law. *Id.* (Kurtzberg II Decl. Ex. 5). On May 23, 2025, the court entered a consent judgment under which the defendants agreed that the robocalls may have violated the Voting Rights Act, and to "not intimidate, threaten, or coerce voters in connection with U.S. elections" going forward. *Id.* (Kurtzberg II Decl. Ex. 6).

for example, not a single example of harm actually caused by "materially deceptive content" that would be addressed by the statute but not by defamation or other existing law. Nor is there even a hypothetical example of such content provided. That is insufficient for strict scrutiny.

II. AB 2655 Violates and is Preempted by 47 U.S.C. §§ 230(c)(1) and 230(c)(2).

AB 2655 goes to Section 230's core: it imposes liability on covered platforms for content created by third parties. Defendants' two efforts to evade preemption under Section 230 fail.

First, Defendants argue—with little more than an inapposite citation to *Black's Law Dictionary*—that Section 230 is inapplicable because AB 2655 does not authorize suits in which the plaintiff can "recover monetary damages against a platform." Defs.' Br. at 28. But courts across the country, including this Court, have rejected Defendants' argument, holding that Section 230 immunizes interactive computer service providers from claims for injunctive relief concerning content posted by third parties on their platforms. Those courts have recognized that to rule otherwise would undermine the entire purpose of Section 230 immunity.

For instance, in *Republican Nat'l Comm.* v. *Google, Inc.*, 2023 WL 5487311, at *8 (E.D. Cal. Aug. 24, 2023), Judge Calabretta explicitly "reject[ed]" the same argument Defendants advance here—that a "claim for injunctive relief is not barred by section 230." Judge Calabretta explained that "the burden of complying with a potential injunction and the burden of litigation itself" would (1) "thwart the stated goal of section 230 'to remove disincentives for the development and utilization of blocking and filtering technologies,'" *id.* (quoting 47 U.S.C. § 230(b)(4)), and (2) "undermine the Ninth Circuit's characterization of section 230 as an immunity *from suit*," *id.* (citing *Fair Hous. Council of San Fernando Valley* v. *Roommates.Com, LLC*, 521 F.3d 1157, 1175 (9th Cir. 2008) (emphasis in original)).

Defendants' position is also squarely contradicted by Section 230's plain text, which confirms that it protects interactive computer service providers from "liability." *See* 47 U.S.C. § 230(e)(3). To interpret "liability" as referring only to monetary payment is too cramped a reading and could lead to the perverse result where Section 230 would prevent providers from paying monetary fines but not federal criminal liability, including imprisonment (*see* § 230(e)(1)).

Defendants' position likewise conflicts with AB 2655's legislative history, wherein the California Legislature recognized that "the California Supreme Court" has "found Section 230 immunity can extend to liability for solely injunctive relief." Kurtzberg Decl. Ex. 4 (Doc. 47-5) at 14 of 26 (citing *Hassell* v. *Bird*, 5 Cal. 5th 522, 547 (2018)); *see also id.* ("Courts have applied Section 230 in a vast range of cases to immunize internet platforms from 'virtually all suits arising from third-party content.'") (quoting Jeff Kosseff, *The Twenty-Six Words That Created The Internet* 57–65 (2019), and citing, *e.g.*, *Carfano* v. *Metrosplash.com, Inc.*, 339 F.3d 1119, 1125 (9th Cir. 2003)).

Indeed, courts across the country have uniformly held that Section 230's protections apply to suits for injunctive relief. *See, e.g.*, *Asia Econ. Inst.* v. *Xcentric Ventures LLC*, 2011 WL 2469822, at *7 (C.D. Cal. May 4, 2011) (holding that "injunctive relief" was "unavailable" under Section 230's "expansive reading that Congress intended"); *Chukwurah* v. *Google, LLC*, 2020 WL 510158, at *2 (D. Md. Jan. 31, 2020) (Section 230 immunity "includes requests for injunctive relief requiring removal or altering disputed third-party content."); *Wells* v. *Youtube, LLC*, 2021 WL 2652966, at *4 (N.D. Tex. May 17, 2021) ("The expansive scope of CDA immunity has been found to encompass . . . requests for injunctive relief."); *Lewis* v. *Google, Inc*., 2021 WL 211495, at *3 (W.D. Pa. Jan. 21, 2021) ("§ 230 immunity has been found to bar injunctive relief."); *Roca Labs, Inc.* v. *Consumer Op. Corp*., 140 F. Supp. 3d 1311, 1319 (M.D. Fla. 2015) ("Section 230 also provides immunity from injunctive and declaratory relief."); *Hinton* v. *Amazon.com.dedc, LLC*, 72 F. Supp. 3d 685, 689 (S.D. Miss. 2014) ("The expansive scope of CDA immunity has been found to encompass . . . requests for injunctive relief."); *Smith* v. *Intercosmos Media Grp., Inc*., 2002 WL 31844907, at *4–5 (E.D. La. Dec. 17, 2002) ("any claim made by the plaintiffs for damages or injunctive relief" is "precluded by the immunity afforded by Section 230(c)(1), and subject to dismissal"); *Noah* v. *AOL Time Warner, Inc.*, 261 F. Supp. 2d 532, 540 (E.D. Va. 2003) ("§ 230 should not be read to permit claims that request only injunctive relief").

Here, Sections 20515(b) and 20516 explicitly authorize suits for "injunctive or other equitable relief," including to "compel the removal of specific content" and the "labeling of specific content." Because such suits interfere with the platforms' role as a publisher—*see, e.g.*, *Barnes* v. *Yahoo!, Inc.*, 570 F.3d 1096, 1102 (9th Cir. 2009) (role of "publisher" includes "reviewing, editing,

and deciding whether to publish or withdraw from publication third-party content"); *Calise* v. *Meta Platforms, Inc.*, 103 F.4th 732, 744 (9th Cir. 2024) (Section 230 precludes liability for "fail[ing] to remove" content); *Hassell*, 5 Cal. 5th at 544–45 (Section 230 barred "cause[s] of action" directing Yelp to remove defamatory consumer reviews)—AB 2655 is preempted by Section 230(c)(1).

Second, this Court should reject Defendants' novel argument that, in the wake of *Moody*, a platform's decision-making about content is not immunized by Section 230. Defendants ask this Court to adopt the reasoning of *Anderson* v. *TikTok, Inc.*, 116 F.4th 180 (3d Cir. 2024), and thus to overturn longstanding Ninth Circuit precedent about Section 230. This Court should decline to do so. In any event, *Anderson*, which was decided based on materially dissimilar facts—i.e., it involved a cause of action arising out of the particular characteristics of TikTok's "expressive algorithm" (*id.* at 184)—does not control here.

Dyroff v. *Ultimate Software Grp., Inc.*, 934 F.3d 1093 (9th Cir. 2019) squarely rejects the argument that, if an interactive computer service provider makes editorial decisions about particular third-party content, the provider loses Section 230 immunity for liability deriving from the content. In *Dyroff*, the plaintiff (an individual who overdosed from drugs he bought on a website) sought to hold the defendant (the website's operator) liable, on the theory that, by using "recommendation and notification functions" to steer the plaintiff to groups on the site facilitating sales of narcotics, the operator became an "information content provider" to which Section 230 did not apply. *Id.* at 1096. *Dryoff* explained that a website operator does not become "an information content provider, losing its Section 230 immunity, by facilitating communication" through "website functions like group recommendations and post notifications." *Id.* at 1097. Although such tools "facilitate the communication and content of others," they "are not content in and of themselves." *Id.* at 1098. Defendants' argument—that AB 2655 "interferes with the[] ability" of covered platforms "to exercise editorial discretion regarding what content to show a user," and thereby only "regulates the speaker's own expressive activity," Defs.' Br. at 31—cannot be reconciled with *Dyroff*.

Numerous courts have held that nothing in *Moody* changes the *Dryoff* analysis. Most recently, *Doe (K.B.)* v. *Backpage.com, LLC*, 2025 WL 719080 (N.D. Cal. Mar. 3, 2025) addressed whether and to what extent *Moody* impacts the protections afforded by Section 230 and rejected

the very argument Defendants advance here. In *Doe*, the plaintiff asserted federal sex trafficking claims against Meta. *Id.* at *1. She alleged that "Meta's algorithms facilitated her sex trafficking by [c]reating, suggesting, and encouraging connections between sex traffickers and vulnerable persons on Instagram and [r]ecommending that vulnerable persons become Instagram 'friends' with sex traffickers." *Id.* at *3. The court understood the plaintiff's argument to be that, because "Meta's algorithms are themselves Meta's content," Section 230 does not apply. *See id.* The plaintiff in *Doe* argued, relying on *Anderson*, that *Moody* overruled *Dyroff*. *Id.* at *4. Judge Lin conducted a thoughtful analysis of this question and rejected the plaintiff's argument. The court held that it did "not read *NetChoice* as overruling *Dyroff*" and that "*NetChoice* does not address Section 230 liability." *Id.* The court explained that the plaintiff's argument:

> appears to presuppose that editorial decisions cannot be both an expression of a publisher's point of view (protected under the First Amendment) and a publication of a third-party's content (protected under Section 230). However, Doe provides ***no basis on which the Court should conclude that Section 230 immunity is mutually exclusive with First Amendment protection***. Indeed, the undisputed core of Section 230 immunity protects a website's moderation decisions, in its role as a 'publisher,' about which third-party content to remove and which to permit. ***That those moderation decisions are also protected by the First Amendment does not strip them of their Section 230 immunity. To hold otherwise would effectively render the core of Section 230 a nullity, contrary to Congress's intent and the plain language of the statute.*** *Id.*

Other courts have likewise repudiated Defendants' reading of *Anderson* on this point. *See M.P. by & through Pinckney* v. *Meta Platforms Inc.*, 127 F.4th 516, 524–26 (4th Cir. 2025) (finding that Section 230 barred tort claims that "solely deal[t] with Facebook's algorithm," because Facebook did not lose Section 230 immunity by choosing to "automate much of its editorial decision-making"); *id.* at 530–31 (Rushing, J., concurring in part and dissenting in part) (explicitly acknowledging *Anderson* and nonetheless rejecting attempt to evade Section 230 by "characteriz[ing] the curated collection of negative third-party content Facebook displays on a susceptible user's News Feed as Facebook's own speech"); *Doe* v. *WebGroup Czech Republic, a.s.*, 2025 WL 879562, at *6 (C.D. Cal. Feb. 21, 2025) ("*Moody* and subsequent circuit decisions like *Anderson* do not disturb the long-standing precedent that online service providers do not create content or lose Section 230 immunity simply by implementing content-neutral algorithms that generate related searches or user-uploaded content based on the users' own viewing activity.").

The factually distinct *Anderson* cannot control here. In that case, the Third Circuit's holding was limited to "curat[ing] compilations of others' content via [] expressive algorithms." 116 F.4th at 184 (citing *Moody*, 603 U.S. at 744). The *Anderson* court held that Section 230 did not immunize TikTok from liability arising from its "'For You Page' ('FYP')" "algorithm [that] curates and recommends a tailored compilation of [content] based on a variety of factors, including the user's age and other demographics, online interactions, and other metadata." *Id.* at 182. Relying on *Moody*'s holding that "a platform's algorithm that reflects editorial judgments about compiling the third-party speech it wants in the way it wants is the platform's own expressive product and is therefore protected by the First Amendment," *Anderson* concluded that TikTok's "FYP algorithm," which "[d]ecid[es] on the third-party speech that will be included in or excluded from a compilation—and then organiz[es] and present[s] the included items on users' FYPs," was TikTok's "own expressive activity." *Id.* at 184.

But here, liability for covered platforms under the statute derives from the characteristics of the content posted by third parties, rather than from the platform's algorithm. *Anderson* itself recognized that Section 230 "immunize[s]" social media platforms to the extent that they are "sued for someone else's expressive activity or content (i.e., third-party speech)." *Id.* at 183. That is precisely what AB 2655 does: it makes it unlawful for platforms to host certain election-related third-party content. Accordingly, *Anderson* is not only distinguishable but also supports "immuniz[ing]" social media platforms from suits under AB 2655.

As the *Backpage* court concluded, Defendants' argument, if accepted, proves too much because it would effectively eliminate ***all*** immunity under Section 230. If, under *Moody*, editorial judgments about content created by third parties can be treated as publishing decisions of the covered platforms, then the platforms will ***always*** be treated as the publisher of content created by others. Under Defendants' reading of the statute, even the classic case for Section 230 immunity—a third party's defamatory post on a website that is not removed by an interactive computer service provider—would result in no immunity under Section 230(c)(1). That cannot be reconciled with (1) Ninth Circuit case law (*see, e.g.*, *Calise*, 103 F.4th at 744 (Section 230 protects from liability for "fail[ing] to remove" content)); (2) the text of Section 230(c)(1) ("[n]o 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"); or (3) Section 230's legislative history (*see Calise*, 103 F.4th at 739 (explaining that Section 230 was enacted to "encourage internet companies to monitor and remove offensive content without fear that they would thereby becom[e] liable for all defamatory or otherwise unlawful messages that they didn't edit or delete")).

If the Court accepted Defendants' position, interactive computer service providers would be heavily incentivized to censor content to avoid the risk of being liable for users' speech, which is antithetical to the reasons Congress enacted Section 230 in the first place.

Finally, although X Corp.'s Complaint seeks a ruling that Section 230(c)(1) and Section 230(c)(2) preempt AB 2655 (Doc. 38 at 59–60, 63), Defendants moved for summary judgment only as to Section 230(c)(1). Defs.' Br. at 28–32. Summary judgment should be granted to Plaintiffs as to Section 230(c)(2) for the reasons outlined in Plaintiffs' Opening Brief (at 5–6, 38).

III. AB 2655 Violates the First and Fourteenth Amendments to the U.S. Constitution on Overbreadth and Vagueness Grounds.

AB 2655 suffers from fundamental vagueness problems that render its expansive requirements next-to-impossible to understand and comply with. These problems are particularly acute from a First Amendment perspective, because the statutory scheme grants the government and other politically motivated individuals unfettered discretion to bring suits for equitable relief that, even if ultimately unsuccessful, will heavily incentivize the censorship of speech. The statute is therefore unconstitutionally overbroad and vague.

Under the First Amendment, "even a law with a plainly legitimate sweep may be struck down in its entirety" as overbroad if the "law's unconstitutional applications substantially outweigh its constitutional ones." *Moody*, 603 U.S. at 723–24. As discussed above, AB 2655's unconstitutional applications substantially outweigh any constitutional ones. *Supra*, Sec. II(b). AB 2655 requires large online platforms to remove or label protected images of then-candidate Trump being arrested. It prohibits Kohls's video. And it demands that platforms remove or label any content presenting even a close call under the statute.

As to vagueness, California argues for a less stringent test than the First Amendment

requires. *See* Defs.' Br. at 25 (reciting test for whether a "statute is unconstitutionally vague in violation of the Due Process Clause"); *id.* at 28 (asserting that AB 2655 "survive[s] scrutiny under the Fourteenth Amendment"); *but see Minn. Voters All.* v. *Mansky*, 585 U.S. 1, 21 (2018) (finding that an "indeterminate prohibition" that carries the "opportunity for abuse" fails First Amendment scrutiny); *Village of Hoffman Ests.* v. *Flipside, Hoffman Ests., Inc.*, 455 U.S. 489, 499 (1982) (applying "stringent vagueness test" to laws that "interfere[] with the right of free speech"); *NAACP* v. *Button*, 371 U.S. 415, 432 (1963) (holding that the "standards of permissible statutory vagueness are strict in the area of free expression"). AB 2655 flunks that strict standard because it fails to provide "a reasonable opportunity to know what" the statute "prohibit[s]." *Hunt* v. *City of Los Angeles*, 638 F.3d 703, 712 (9th Cir. 2011).

First, the statute's definition of "materially deceptive content" is unconstitutionally vague. *See* §§ 20512(i), 20513, 20514. The definition—content that would "falsely appear to a reasonable person to be an authentic record of the content depicted in the media," § 20512(i)(1)—forces covered platforms and content creators to "necessarily guess at" what types of digitally altered content would appear authentic enough to a hypothetical "reasonable person" to determine if they are covered by the law. *FCC* v. *Fox Television Stations, Inc.*, 567 U.S. 239, 253 (2012); *see also, e.g.*, *Tucson Woman's Clinic* v. *Eden*, 379 F.3d 531, 555 (9th Cir. 2004) (striking down a regulation on vagueness grounds that "subjected physicians to sanctions based not on their own objective behavior, but on the subjective viewpoint of others"), *abrogated on other grounds*, *Dobbs* v. *Jackson Women's Health Org.*, 597 U.S. 215 (2022).

Determining whether content is "materially deceptive" is unworkable in practice, as the factual record makes clear. *See* Pls.' Br. at 40; PSUF ¶ 182 (citing X Corp. S&O Decl. ¶¶ 6, 19; Rumble Decl. ¶ 45) (discussing the burden of making these subjective "judgment calls" on X's and Rumble's manual content reviewers); *see also id.* ¶¶ 33–35 (citing Dillon Decl. ¶¶ 58–64) (showing that Snopes thought it was necessary to "fact-check" a satirical post). Defendants attempt to defend the definition of "materially deceptive content" as "not unduly vague," by again erroneously asserting that it "tracks the line drawn under the First Amendment for when parody or satire can be actionable as defamation." Defs.' Br. at 26. But this argument fails because, as discussed above

and as this Court found, the statute "does ***much more*** than punish potential defamatory statements." *Kohls*, 752 F. Supp. at 1193. Defendants even concede that the statute covers more than defamatory speech by stating that "existing causes of action, such as defamation law, do not adequately address the harms that AB 2655 seeks to prevent[.]" Defs.' Br. at 23.

Second, AB 2655's vagueness and the difficulties of complying with its uncertain requirements are compounded because, not only must the covered platforms and content creators determine whether content is "materially deceptive" under Section 20512(i), but the Blocking and Labeling Requirements also force them to determine whether the content is "reasonably likely to harm the reputation or electoral prospects of a candidate," and/or "reasonably likely to undermine confidence in the outcome of one or more election contests." §§ 20513(a)(2), 20514(a)(2). As this Court found, the term "electoral prospects" is "amorphous" and "difficult to ascertain," and the inquiry of whether content is reasonably likely to harm an individual's "electoral prospects" "lack[s] any objective metric." *Kohls*, 752 F. Supp. 3d at 1193–94. Even worse, these judgments need to be made with the knowledge that, in any close case, a lawsuit might be brought against any covered platform if questionable content is not taken down or labeled, and that no such lawsuit can be brought under AB 2655 if material *is* taken down or labeled.

Third, in addition to making these highly subjective judgment calls under Sections 20512(i), 201513(a)(2), and 20514(a)(2), covered platforms and creators must also determine whether the content is "satire or parody," such that it would be exempted under Section 20519. But as discussed, this too is a highly subjective inquiry about which reasonable minds will often disagree. And the record demonstrates this, because although Plaintiff Kohls included the word "parody" in the "Kamala Harris Ad PARODY" video title, Governor Newsom indicated that, in his view, the video would be prohibited by AB 2655. *Supra*, Sec. I(a). Nor was Governor Newsom dissuaded about the statute's applicability to the "Kamala Harris Ad PARODY" by the fact that, in the video, "Harris" states that she is a "diversity hire" and "deep state puppet" who "may not know the first thing about running the country." *See* Pls.' Br. at 18. And Defendants now emphasize that the video was reposted without the parody disclaimer. DSUF ¶ 3. Because AB 2655 provides "no principle for determining when" speech will "pass from the safe harbor" to "the forbidden," it is

unconstitutionally vague. *Gentile* v. *State Bar of Nev.*, 501 U.S. 1030, 1049 (1991).

Fourth, the State attempts to defend the Blocking Requirement by repeating its faulty argument that it does not compel covered platforms to take any "affirmative action." Defs.' Br. at 27. That is inconsistent with the statute's plain language, as set forth above. *See supra*, Sec. II(b) (explaining that AB 2655 explicitly requires platforms to affirmatively "identify" and "remove" certain content). Even worse, the platforms must satisfy this affirmative obligation by using "state-of-the-art techniques" (a term that the statute does not define). § 20513(c). Perhaps unsurprisingly, Defendants' brief is entirely silent on the statute's "state-of-the-art techniques" requirement.

Fifth, the carve-out within the definition of "materially deceptive content" for "audio or visual media that contains only minor modifications that do not significantly change the perceived contents or meaning of the content," is also vague. § 20512(i)(2). Determining whether modifications to particular content are "***minor***" or "significant" is likewise a highly subjective inquiry, as the "boundary between a 'significant' and an insignificant nexus is far from clear." *Sackett* v. *EPA*, 598 U.S. 651, 681 (2023).

The result of these ambiguities—compounded with (1) the statute's unreasonable timeframes, (2) the amount of content for which platforms will need to make these determinations, and (3) the fact that the statute entirely immunizes over-censoring and over-labeling content, Pls.' Br. at 41–42; *id.* at 12–20—is exactly the sort of vague statutory scheme, ripe for "resolution on an *ad hoc* and subjective basis, with arbitrary and discriminatory application," that the Constitution prohibits. *Hunt*, 638 F.3d at 712; *see also, e.g.*, *Button*, 371 U.S. at 432–33 ("The objectionable quality of vagueness and overbreadth" in "the area of First Amendment freedoms" is "the existence of a penal statute susceptible of sweeping and improper application.") And here, in the context of elections, those concerns are particularly severe, because the politically ***driven*** enforcers of the statute can apply it to particular content and viewpoints to further their own political agendas.

CONCLUSION

The Court should deny Defendants' motion for summary judgment, grant Plaintiffs' motion for summary judgment, declare AB 2655 unconstitutional and illegal, and permanently enjoin its enforcement.

Dated: June 6, 2025

By: /s/ Joel Kurtzberg
CAHILL GORDON & REINDEL LLP
Joel Kurtzberg (admitted *pro hac vice*)
Floyd Abrams (admitted *pro hac vice*)
Jason Rozbruch (admitted *pro hac vice*)
32 Old Slip
New York, NY 10005
Phone: 212-701-3120
Facsimile: 212-269-5420
jkurtzberg@cahill.com

DOWNEY BRAND LLP
William R. Warne (SBN 141280)
Meghan M. Baker (SBN 243765)
621 Capitol Mall, 18th Floor
Sacramento, CA 95814
Phone: 916-444-1000
Facsimile: 916-520-5910

Attorneys for Plaintiff X Corp.

/s/ Adam E. Schulman
HAMILTON LINCOLN LAW INSTITUTE
Adam E. Schulman (admitted *pro hac vice*)
Theodore H. Frank (SBN 196332)
1629 K Street NW, Suite 300
Washington, DC 20006
Phone: 610-457-0856
adam.schulman@hlli.org
ted.frank@hlli.org

Attorneys for Plaintiff Christopher Kohls

/s/Johannes Widmalm-Delphonse
ALLIANCE DEFENDING FREEDOM
Johannes Widmalm-Delphonse
44180 Riverside Parkway
Lansdowne, VA 20176
Phone: 571-707-4655
jwidmalmdelphonse@ADFlegal.org

ALLIANCE DEFENDING FREEDOM
Jonathan A. Scruggs
15100 N. 90th Street
Scottsdale, AZ 85260
Phone: 480-444-0020
Facsimile: 480-444-0028
jscruggs@ADFlegal.org

CHAVEZ-OCHOA LAW OFFICES, INC.
Brian R. Chavez-Ochoa (SBN 190289)
4 Jean Street, Suite 4
Valley Springs, CA 95252
Phone: 209-772-3013
brianr@chavezochoalaw.com

Attorneys for Plaintiffs The Babylon Bee, LLC and Kelly Chang Rickert

/s/ Mathew W. Hoffmann
ALLIANCE DEFENDING FREEDOM
Mathew W. Hoffmann
Philip A. Sechler
44180 Riverside Parkway
Lansdowne, VA 20176
Phone: 571-707-4655
Facsimile: 571-707-4656
mhoffmann@ADFlegal.org
psechler@ADFlegal.org

ALLIANCE DEFENDING FREEDOM
Mercer Martin
15100 N. 90th Street
Scottsdale, AZ 85260
Phone: 480-444-0020
Facsimile: 480-444-0028
mmartin@ADFlegal.org

CHAVEZ-OCHOA LAW OFFICES, INC.
Brian R. Chavez-Ochoa (SBN 190289)
4 Jean Street, Suite 4
Valley Springs, CA 95252
Phone: 209-772-3013
brianr@chavezochoalaw.com

Attorneys for Plaintiffs Rumble Inc. and Rumble Canada Inc.

PROOF OF SERVICE

I hereby certify that I caused to be filed a true and accurate copy of the foregoing document with the Clerk of Court using the CM/ECF system, which automatically sends an electronic notification to all attorneys of record in this case.

Dated: June 6, 2025

By: /s/ Joel Kurtzberg
CAHILL GORDON & REINDEL LLP
Joel Kurtzberg (admitted *pro hac vice*)
32 Old Slip
New York, NY 10005
Phone: 212-701-3120
Facsimile: 212-269-5420
jkurtzberg@cahill.com