

# Beyond Coercion: Rethinking State Action Doctrine for Government-Platform Speech Regulation

Samantha Murray

## Abstract

*When government officials pressure social media platforms to remove content, does the First Amendment apply? Current state action doctrine answers this question with a binary test: either the government coerced the platform, triggering constitutional scrutiny, or it merely persuaded, leaving users without recourse. The Supreme Court's recent decision in *Murthy v. Missouri* exposed the inadequacy of this framework. Faced with a sustained campaign of government pressure combining repeated high-level contacts, veiled regulatory threats, and positive reinforcement for compliance, the Court could not reach the constitutional merits and was dismissed on standing grounds. This Article argues that the coercion-persuasion binary fails to capture the spectrum of modern government-platform relationships and proposes a sliding scale framework as an alternative. Rather than asking whether the government coerced the platform in some binary sense, courts should weigh the degree of government pressure against the strength of the platform's independent editorial judgment. This approach accounts for platforms' own First Amendment rights, recognized in *Moody v. NetChoice*, while protecting users from censorship that operates through private intermediaries influenced by the state. The framework preserves legitimate government-platform cooperation while identifying the point at which sustained pressure transforms nominally private moderation decisions into unconstitutional state action.*

## I. Introduction

When the White House contacts Facebook to express concern about vaccine misinformation, when the FBI alerts Twitter to foreign disinformation campaigns, when the Surgeon General publicly calls on platforms to combat health misinformation, are these legitimate communications between government and private actors, or are they unconstitutional censorship by proxy? The answer matters profoundly for the future of online speech, yet current First Amendment doctrine provides no clear guidance.<sup>1</sup>

The Supreme Court's recent decision in *Murthy v. Missouri* illustrates the problem. The case involved extensive government engagement with social media platforms regarding COVID-19 misinformation and election-related content. White House officials, the CDC, and the FBI held repeated meetings with platforms, flagged specific posts and accounts for removal, expressed frustration when platforms responded slowly, made veiled references to potential Section 230

---

<sup>1</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 1972 (2024).

reform,<sup>2</sup> and praised platforms that complied swiftly. Platforms subsequently removed substantial amounts of content in these areas.<sup>3</sup> Yet the Supreme Court never reached the constitutional question, instead dismissing the case on standing grounds. The Court held that plaintiffs failed to prove their speech was suppressed because of government pressure rather than platforms' independent content-moderation policies.<sup>4</sup>

This disposition reveals a fundamental inadequacy in state action doctrine. The traditional framework treats government influence as a binary question: either the government coerced a private party—triggering constitutional scrutiny—or it merely persuaded—causing no constitutional concern. This binary test made sense in simpler contexts. In *Bantam Books, Inc. v. Sullivan*, Rhode Island officials either threatened book distributors with prosecution or they did not.<sup>5</sup> The choice was clear-cut. But modern government-platform relationships operate in a vast middle ground between pure coercion and pure persuasion. Officials engage in sustained communication campaigns combining implicit threats, explicit criticism, positive reinforcement, and appeals to shared interests. Platforms respond by modifying their content moderation in ways that might reflect both government pressure and their own editorial judgment. The binary framework cannot capture this complexity.

The problem is compounded by platforms' own First Amendment rights. As the Supreme Court recognized in *Moody v. NetChoice* and *NetChoice v. Paxton*, social media platforms exercise protected editorial discretion when they moderate content.<sup>6</sup> Just as the government cannot force a newspaper to publish content it wishes to exclude, it cannot force Facebook or YouTube to host speech those platforms wish to remove. Platforms' editorial autonomy is itself a constitutional value.<sup>7</sup> Yet this creates a puzzle: if platforms have a right to moderate content, how can their moderation decisions ever constitute state action subject to First Amendment scrutiny? When government officials communicate with platforms about content and platforms subsequently remove that content, courts must determine whether the platform exercised independent editorial judgment—which is protected—or acted under government pressure—which can be potentially unconstitutional. The binary coercion test provides no mechanism for making this determination when, as in *Murthy*, government influence and platform judgment coexist.

The result is a three-way constitutional tension that current doctrine cannot resolve. The government has legitimate interests in communicating with platforms about matters of public concern—foreign interference, public health emergencies, and violent threats. Platforms have First Amendment rights to curate content on their services according to their values and business interests. And users have First Amendment rights to speak without government censorship, even

---

<sup>2</sup> Id.

<sup>3</sup> Id.

<sup>4</sup> Id.

<sup>5</sup> *Bantam Books, Inc. v. Sullivan*, 372 U.S. 58 (1963).

<sup>6</sup> *Moody v. NetChoice, LLC*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024); *NetChoice, LLC v. Paxton*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>7</sup> U.S. CONST. amend. I.

when that censorship operates through private intermediaries.<sup>8</sup> Existing state action doctrine privileges the first two interests at the expense of the third. By setting the bar for "coercion" so high that it rarely applies when platforms have any independent interest in content moderation, the doctrine effectively allows the government to launder censorship through platforms' editorial discretion.

This Article proposes a new framework for analyzing government influence over platform content moderation. Rather than asking whether the government coerced the platform in some binary sense, courts should apply a sliding scale that weighs the degree of government pressure against the strength of the platform's independent editorial judgment. The more coercive the government's conduct and the less it aligns with what the platform would independently do, the stronger the First Amendment problem. This approach recognizes that state action exists on a continuum and provides courts with tools to distinguish legitimate government communication from unconstitutional pressure.

## II. The Binary Framework and Its Limits

The state action doctrine has long policed the boundary between permissible government persuasion and unconstitutional coercion of private actors. Courts apply a binary test: if the government coerces a private party to restrict speech, that private restriction becomes state action subject to First Amendment scrutiny. If the government merely persuades or encourages, no constitutional violation occurs, even if the private party ultimately restricts speech the government dislikes.<sup>9</sup>

The binary framework made sense in simpler contexts. In *Bantam Books, Inc. v. Sullivan*, the government's threats of criminal prosecution and regulatory sanctions clearly crossed into coercion. The choice was binary: either Rhode Island's Commission to Encourage Morality in Youth wielded coercive power over book distributors or it did not. The Court found coercion and invalidated the scheme as an unconstitutional prior restraint.<sup>10</sup> But *Murthy v. Missouri* exposed the inadequacy of this binary approach in the context of government engagement with major platforms regarding COVID-19 misinformation and election-related content. White House officials, the CDC, and the FBI held repeated meetings with platforms, flagged specific posts and accounts for removal, and expressed frustration when platforms failed to comply quickly enough. In one notable exchange, a White House official asked Meta representatives, "Are you guys f'ing serious? I want an answer on what happened here and I want it today."<sup>11</sup>

The Supreme Court has distinguished between these categories for decades. In *Blum v. Yaretsky*, the Court held that a state "normally can be held responsible for a private decision only when it has exercised coercive power or has provided such significant encouragement, either overt or

---

<sup>8</sup> *Blum v. Yaretsky*, 457 U.S. 991, 1004 (1982).

<sup>9</sup> *Bantam Books*, 372 U.S. at 67.

<sup>10</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>11</sup> *Id.*

covert, that the choice must in law be deemed to be that of the State."<sup>12</sup> The framework asks

whether the government's involvement was so pervasive that the private actor's decision loses its private character and becomes attributable to the state itself.<sup>13</sup>

This test works well at the extremes. Explicit threats of prosecution, loss of licenses, or statutory sanctions clearly constitute coercion.<sup>14</sup> Conversely, public statements of government preference or requests for voluntary cooperation fall comfortably within permissible persuasion. The government may use its bully pulpit to advocate for private action without triggering First Amendment scrutiny.<sup>15</sup> The binary framework reflects an important constitutional value: private actors retain their own constitutional rights, including editorial discretion over their platforms. When Facebook or Twitter removes content based on their own judgment, even judgment informed by government input, they exercise their First Amendment rights, not government power. The state action doctrine must therefore identify when government involvement crosses the line from influencing a genuinely private decision to effectively commandeering private editorial authority.

But what happens when government pressure operates in the vast gray area between clear coercion and pure persuasion? When the most powerful officials in the federal government repeatedly contact platforms, express displeasure with their moderation decisions, and hint at regulatory consequences—but stop short of explicit threats? The binary framework struggles to capture this reality, as *Murthy v. Missouri* demonstrates.<sup>16</sup>

### **III. Murthy v. Missouri and the Failure of Binary Analysis**

*Murthy v. Missouri* arose from an extensive campaign of government engagement with social media platforms regarding COVID-19 misinformation and election-related content. Beginning in 2021, officials from the White House, the CDC, the Surgeon General's office, and the FBI engaged in regular communication with Facebook, Twitter, and YouTube about content moderation. The government's conduct combined elements that, viewed in isolation, might appear benign. Officials held meetings to discuss misinformation trends. They flagged specific posts and accounts they believed spread false information. They requested data on platform enforcement actions. Standing alone, such engagement could plausibly constitute information-sharing or public health advocacy—permissible persuasion rather than unconstitutional coercion.<sup>17</sup>

But the pattern of conduct revealed a more troubling dynamic. When platforms failed to act on flagged content quickly enough, officials expressed pointed frustration. White House staff

---

<sup>12</sup> Id.

<sup>13</sup> Blum, 457 U.S. at 1004.

<sup>14</sup> Id.

<sup>15</sup> Id.

<sup>16</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>17</sup> Communications Decency Act of 1996, 47 U.S.C. § 230 (2018).

publicly criticized platform efforts as inadequate. More significantly, government officials repeatedly made references to potential regulatory action. These references to Section 230

reform,<sup>18</sup> the legal provision that shields platforms from liability for user content, hung over every interaction. The implicit message was clear: cooperate with our content moderation preferences, or face legislative retaliation.<sup>19</sup>

Conversely, when platforms complied swiftly with government requests, officials offered effusive praise. They thanked platforms for their partnership, commended their responsiveness, and reinforced that such cooperation reflected the appropriate relationship between government and platforms. This positive reinforcement revealed that the government was not merely sharing information or expressing viewpoints—it was actively attempting to shape platform behavior through a combination of implicit threats and rewards. The result was measurable: platforms removed or suppressed substantial amounts of content after government contact, including content they had not previously identified as violating their own policies.<sup>20</sup>

Despite this pattern, the Supreme Court in *Murthy v. Missouri* never reached the question of whether this conduct constituted unconstitutional coercion. Instead, the 6-3 majority dismissed the case on standing grounds, holding that plaintiffs failed to prove their speech was suppressed because of government pressure rather than platforms' independent content moderation policies.<sup>21</sup> The Court noted that "the platforms moderated similar content long before any of the Government defendants engaged in" the challenged conduct.<sup>22</sup>

This disposition reveals as much as it conceals. The Court's inability to reach the merits stems from a deeper problem: the binary coercion framework lacks the tools to analyze government conduct that operates through sustained influence rather than explicit threats. When platforms have preexisting content moderation policies that align somewhat with government preferences, how can courts determine whether subsequent enforcement of those policies results from the platforms' independent judgment or government pressure? The binary test offers no mechanism to answer this question. It simply asks whether the government crossed a threshold into coercion, without accounting for how government involvement might amplify, accelerate, or alter enforcement of platform policies in ways that burden speech.<sup>23</sup>

The *Murthy* Court's standing dismissal sidesteps a fundamental inadequacy in state action doctrine: the coercion versus persuasion binary cannot capture government influence that operates through sustained pressure on platforms with preexisting moderation incentives. The framework treats state action as an on-off switch—either the government coerced the platform

---

<sup>18</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>19</sup> *Id.*

<sup>20</sup> *Id.*

<sup>21</sup> *Id.*

<sup>22</sup> *Id.*

<sup>23</sup> *Id.*

(state action) or it merely persuaded (no state action). But *Murthy* reveals a spectrum of government conduct that resists this categorization.<sup>24</sup>

Consider the government's approach in *Murthy*. Officials did not issue explicit threats of legal action, as in *Bantam Books*. They did not mandate specific outcomes through regulation. Instead, they employed a sophisticated strategy combining public criticism, private pressure, veiled regulatory threats, and positive reinforcement for compliance. This conduct was simultaneously too aggressive to dismiss as mere persuasion, yet too subtle to constitute clear coercion under existing doctrine.

The binary framework fails in three distinct ways. First, it ignores degrees of pressure. Not all government communications carry equal weight. A casual suggestion from a mid-level bureaucrat differs fundamentally from repeated demands by White House officials coupled with public statements criticizing platforms and implicit threats about Section 230 reform.<sup>25</sup> Yet the binary test treats these as the same unless the pressure crosses some threshold into actionable coercion. This all-or-nothing approach creates a massive zone of unregulated government influence. Officials can pressure platforms aggressively, short of explicit threats, without triggering constitutional scrutiny.

Second, the framework neglects platforms' independent editorial judgment. The *Murthy* majority emphasized that "the platforms moderated similar content long before any of the Government defendants engaged in" the challenged conduct.<sup>26</sup> This observation was central to the standing analysis—if platforms were already removing COVID misinformation, how could plaintiffs prove government pressure caused the removals? But this reasoning exposes a crucial gap. Platforms having some preexisting interest in content moderation does not mean government pressure played no role in their decisions.

Consider a platform that sporadically removes certain misinformation but tolerates much of it as protected speech. When government officials begin flagging content, praising swift removals, and hinting at regulatory consequences, the platform might dramatically expand its enforcement, taking down content it previously would have left up. The government has successfully altered the platform's editorial calculus, yet the binary framework offers no way to recognize this influence. Because the platform had some independent interest in the category of speech, the government's role becomes invisible.<sup>27</sup>

This is precisely what makes *Murthy* troubling. That platform may well have removed some COVID misinformation absent government involvement. But did they remove all the content? Did they remove it as quickly or broadly? Did government pressure cause them to err on the side of removal rather than leaving borderline content up? The binary framework cannot answer these

---

<sup>24</sup> Id.

<sup>25</sup> Id.

<sup>26</sup> Communications Decency Act of 1996, 47 U.S.C. § 230 (2018).

<sup>27</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

questions because it only asks whether the government coerced a decision the platform otherwise would not have made—not whether government pressure shaped a decision the platform made partially on its own.<sup>28</sup>

Third, the framework fails to account for users' First Amendment interests. State action doctrine asks only whether the government coerced the platform, treating the platform as the relevant constitutional actor. But when platforms remove speech, users suffer the injury. The constitutional concern is not just that the government pressured a private party—it is that government pressure resulted in speech suppression.<sup>29</sup> Yet the binary test makes users' speech rights entirely contingent on proving that platforms acted solely because of government coercion, not their own judgment. If a platform has any independent reason to moderate content, government pressure disappears from the analysis. This effectively allows the government to launder censorship through platforms' editorial discretion. As long as officials are careful not to issue explicit threats, they can achieve through pressure what the First Amendment would forbid them from doing directly.

The dissent in *Murthy* recognized this problem. Justice Alito wrote that the case represented "one of the most important free speech cases to reach this Court in years" and warned that the majority's decision "permits the successful campaign of coercion in this case to stand as an attractive model for future officials who want to control what the people say, hear, and think."<sup>30</sup> Alito understood that something deeply problematic occurred, even if it did not fit neatly into the binary framework's definition of coercion.

The fundamental flaw is this: the binary framework assumes government involvement either completely determines a private party's decision or is irrelevant. It cannot recognize that government pressure might be one significant factor among several, tipping platforms toward more aggressive content moderation than their independent judgment would support. This conceptual gap leaves a wide swath of government conduct—conduct that meaningfully shapes private censorship—beyond constitutional reach. *Murthy* demonstrates that courts need a framework capable of recognizing government influence that operates along a spectrum, not a binary. The question should not be whether government pressure completely displaced platform judgment, but whether it substantially influenced content moderation decisions in ways that burden protected speech.

#### **IV. Platforms' First Amendment Rights and the Three-Way Tension**

The state action problem in *Murthy* cannot be understood in isolation from platforms' own constitutional rights. Social media companies are not mere conduits for speech—they are editors with First Amendment protection for their content-moderation decisions. This reality, confirmed

---

<sup>28</sup> Id.

<sup>29</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024) [Alito, J., dissenting]

<sup>30</sup> Id.

in *Moody v. NetChoice* and *NetChoice v. Paxton*, fundamentally complicates the analysis of government influence over platform censorship.<sup>31</sup>

The Supreme Court's 2024 *NetChoice* decisions, though ultimately remanded on procedural grounds, established critical principles about platforms' editorial autonomy. Justice Kagan's

majority opinion emphasized that social media platforms exercise "editorial discretion" protected by the First Amendment when they select, curate, and moderate user content.<sup>32</sup> The Court held that "the First Amendment offers protection when an entity engaging in expressive activity, including compiling and curating others' speech, is directed to accommodate messages it would prefer to exclude."<sup>33</sup>

The key precedent was *Miami Herald Publishing Co. v. Tornillo*, in which the Court struck down a Florida law requiring newspapers to give political candidates a right of reply to criticism.<sup>34</sup> Just as Florida could not force the Miami Herald to publish content it wished to exclude, Texas and Florida cannot force Facebook or YouTube to host content those platforms wish to remove.<sup>35</sup> The government's interest in "improving" or "balancing" the marketplace of ideas does not overcome platforms' editorial rights.<sup>36</sup>

This principle creates an immediate tension with state action doctrine in cases like *Murthy*. If platforms have a First Amendment right to remove content they deem harmful, how can their removal decisions ever constitute state action subject to First Amendment scrutiny? After all, when a platform removes misinformation about COVID-19 or election integrity, it exercises its own constitutional rights. The government cannot be held responsible for a platform's independent editorial judgment, even if that judgment results in suppressing protected speech.<sup>37</sup>

But this framing oversimplifies the constitutional landscape. Platforms' editorial rights are not absolute shields against state action scrutiny. The question is not whether platforms can moderate content—of course they can—but whether their moderation decisions remain genuinely independent when made under government pressure.<sup>38</sup>

At one extreme, a platform independently develops content policies, applies them consistently based on its own judgment, and removes content without any government involvement. Here, the platform exercises pure editorial discretion. Users may dislike the platform's choices, but those choices reflect private ordering, not state action. The First Amendment does not guarantee anyone a right to speak on another's private platform.<sup>39</sup> On the other hand, the government

---

<sup>31</sup> *Moody v. NetChoice, LLC*, 144 S. Ct. at 2383; *NetChoice, LLC v. Paxton*, 144 S. Ct. at 2383.

<sup>32</sup> *NetChoice*, 144 S. Ct. at 2395 (majority opinion).

<sup>33</sup> *Id.* at 2396.

<sup>34</sup> *Miami Herald Publ'g Co. v. Tornillo*, 418 U.S. 241, 258 (1974).

<sup>35</sup> *Id.*

<sup>36</sup> U.S. CONST. amend. I.

<sup>37</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>38</sup> *Id.*

<sup>39</sup> *Id.*

explicitly threatens a platform with legal consequences unless it removes specific content. Here, the platform's nominal "choice" to remove content becomes state action.<sup>40</sup> The government has effectively commandeered the platform's editorial function, making the platform an instrument of state censorship. This clearly violates the First Amendment—not the platform's rights, but the users' rights.<sup>41</sup>

In the middle lies the problematic gray area that *Murthy* inhabits. Government officials engage in sustained communication with platforms, expressing strong disapproval of certain content, praising swift removals, and making veiled references to regulatory consequences. Platforms then remove content—but they might have removed some of it anyway based on their preexisting policies. The platforms' editorial rights and government pressure coexist, making it nearly impossible to disentangle independent judgment from governmental influence.<sup>42</sup>

This is where the binary state action framework breaks down. Did the government coerce the platform? But platforms' editorial rights make "coercion" difficult to identify. A platform that has some independent interest in content moderation will claim it acted on its own judgment, not government pressure. The government, for its part, will characterize its communications as mere persuasion or information sharing, not coercion. And because platforms do have genuine editorial interests—they don't want misinformation harming their users or advertisers—it becomes facially plausible that they acted independently.<sup>43</sup>

Yet this cannot be the end of the analysis. Platforms' editorial rights do not give the government license to pressure them into suppressing speech the government dislikes. The fact that a platform could independently choose to remove content does not mean government pressure played no role when it actually does remove that content. Editorial discretion exercised under sustained government pressure is not the same as editorial discretion exercised independently. The constitutional problem is not that platforms moderate content—that is their right. The problem arises when the government influences how they moderate, what they prioritize, and how quickly they act. Even if platforms retain nominal decision-making authority, government pressure can shift their editorial calculus in ways that effectively achieve what the First Amendment forbids the government from doing directly.<sup>44</sup>

This creates a three-way tension that current doctrine cannot resolve. First, the government has legitimate interests in communicating with platforms about matters of public concern. Officials may share information about foreign disinformation campaigns, public health emergencies, or threats to election integrity. Such communication does not automatically become coercion simply

---

<sup>40</sup> See *Bantam Books*, 372 U.S. at 67.

<sup>41</sup> Communications Decency Act of 1996, 47 U.S.C. § 230 (2018).

<sup>42</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>43</sup> *Id.*

<sup>44</sup> *Id.*

because platforms then act on that information.<sup>45</sup> Second, platforms have First Amendment rights to moderate content as they see fit. They need not—and often should not—host certain categories of speech, from graphic violence to deliberate misinformation. Their editorial judgments deserve constitutional protection.<sup>46</sup> Third, users have First Amendment rights to speak without government censorship. When government pressure causes platforms to suppress protected speech that platforms otherwise would have allowed, users suffer a constitutional injury.<sup>47</sup> The fact that a private intermediary carried out the suppression does not cure the government's role in causing it.

Current state action doctrine privileges the first two interests—government speech and platform autonomy—at the expense of the third. By requiring proof of coercion so clear that it overcomes any claim of platform independence, the doctrine makes it nearly impossible to vindicate users' rights when government influence operates subtly. A better framework must account for all three constitutional interests. It should recognize that government communications with platforms exist on a spectrum, that platforms' editorial independence can be genuine or can be compromised by government pressure, and that users' speech interests deserve protection even when censorship occurs through private intermediaries influenced by the state.

## V. A Sliding Scale Framework

The failures of *Murthy* and the tensions revealed by *NetChoice* demand a new approach. Courts need a framework that can recognize government influence operating along a spectrum, account for platforms' varying degrees of editorial independence, and protect users' speech rights without eliminating legitimate government-platform communication. The constitutional concern should increase as government conduct becomes more coercive and platform independence weakens. When both factors are high—substantial government pressure combined with weak platform independence—a constitutional violation exists even without proof of explicit coercion.<sup>48</sup> Conversely, when government involvement is minimal or platform independence is strong, no violation occurs even if speech is suppressed. This approach recognizes that state action exists on a continuum. The question is not whether the government "coerced" in some binary sense, but whether government involvement was sufficiently substantial, relative to the platform's genuine editorial interest, to transform a nominally private decision into state action.

Government pressure exists on a spectrum from low to high. At the low end, the government makes public statements of preference without direct contact with platforms, engages in single instances of information-sharing about threats or trends, makes requests for voluntary cooperation framed as suggestions, or communicates through low-level officials with no regulatory authority. Medium pressure involves repeated communications from senior officials,

---

<sup>45</sup> Id.

<sup>46</sup> Id.

<sup>47</sup> Id.

<sup>48</sup> Id.

expressions of frustration or disappointment with platform responses, public criticism of platforms' moderation efforts, and requests framed with increasing urgency. High pressure includes explicit or implicit threats of regulatory consequences,<sup>49</sup> conditioning government benefits on cooperation, repeated demands for specific actions, praise for compliance combined with criticism for resistance, and communications from officials with direct regulatory authority over platforms.

The *Murthy* conduct fell into the high pressure category. The White House, CDC, and FBI engaged in repeated contacts, expressed frustration when platforms did not act quickly, made veiled references to Section 230 reform,<sup>50</sup> and offered effusive praise when platforms complied. This was not casual information-sharing—it was a sustained campaign to alter platform behavior through a combination of implicit threats and positive reinforcement.<sup>51</sup>

Courts should consider several indicators when assessing pressure. Frequency and persistence matter because a single communication differs fundamentally from dozens of contacts over months. Sustained engagement signals that the government seeks to influence platform behavior, not merely share information. The source of communication is equally important: messages from the White House or agency heads carry more weight than communications from junior officials. The speaker's regulatory authority over platforms matters. Tone and content reveal whether communications are framed as information-sharing or as demands. When officials express frustration, criticism, or disappointment when platforms don't comply, such language reveals an expectation of obedience rather than mere persuasion. Explicit or implicit consequences are critical—when the government references its regulatory power, mentions potential legislation, or conditions cooperation on avoiding adverse action, even veiled threats count as pressure.<sup>52</sup> Finally, public versus private pressure both matter: public criticism of platforms can constitute pressure even without private threats, by creating reputational costs for non-compliance and signaling government disapproval.

The strength of a platform's independent interest in the content moderation at issue is equally critical. This factor asks whether the platform would have taken this action absent government involvement. Strong independence exists when a platform has longstanding, consistently enforced policies against the category of content, initiated the moderation practice before government contact, removes content in this category regardless of government interest, has a stated rationale that aligns with business interests such as user safety or advertiser preferences, and resists government requests that conflict with its policies.<sup>53</sup> Weak independence exists when a platform has no preexisting policy against the content category, only began enforcement after government pressure, enforces policies inconsistently while prioritizing government-flagged

---

<sup>49</sup> Id.

<sup>50</sup> Blum, 457 U.S. at 1004.

<sup>51</sup> U.S. CONST. amend. I.

<sup>52</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>53</sup> Id.

content, shifts its stated rationale to match government preferences, or removes content only when government flags it. Moderate independence exists when a platform has broad policies that could cover the content but rarely enforced them before government contact, expands enforcement dramatically after government involvement, or removes some content in the category independently but accelerates enforcement of government priorities.

Evidence of weak independence might include timing—whether enforcement began or dramatically increased after government contact. Scope matters too: whether the platform removes government-flagged content more quickly or consistently than similar content the government has not mentioned. Inconsistency is relevant when the platform enforces policies more aggressively in areas of government concern than in analogous areas. Stated justifications reveal whether platform explanations mirror government talking points rather than independent business rationales. Internal documents might show platforms acting reluctantly or only due to government pressure.

In *Murthy*, this factor cuts both ways. Platforms had preexisting policies against misinformation and "moderated similar content long before" government involvement.<sup>54</sup> This suggests some independent editorial interest. However, the dramatic increase in enforcement after government contact, the prioritization of government-flagged content, and the speed of response to official communications all suggest government pressure substantially influenced the platforms' editorial calculus.<sup>55</sup> The platforms may have had some independent interest, but government involvement amplified and directed that interest in ways that burden protected speech.

Under this sliding scale, constitutional violations exist when government pressure substantially outweighs platform independence. Several scenarios illustrate the framework's application. Consider first a situation where the FBI informs Facebook of a Russian disinformation campaign targeting U.S. elections, and Facebook—which already prohibits coordinated inauthentic behavior—investigates and removes the accounts. This involves low government pressure combined with strong platform independence, resulting in no constitutional violation. Facebook acts on its own judgment based on government-provided information.<sup>56</sup>

Consider next a situation where the White House repeatedly contacts YouTube demanding removal of vaccine-skeptical content, and YouTube has no policy against vaccine skepticism but creates one after sustained pressure and veiled threats about Section 230 reform.<sup>57</sup> This involves high government pressure combined with weak platform independence—as in no preexisting policy and enforcement only after pressure—resulting in a constitutional violation. The government has effectively commandeered YouTube's editorial function.

A third scenario involves the Surgeon General publicly encouraging platforms to address health

---

<sup>54</sup> Communications Decency Act of 1996, 47 U.S.C. § 230 (2018).

<sup>55</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024)

<sup>56</sup> *Id.*

<sup>57</sup> *Id.*

misinformation, after which Twitter—which already removes some health misinformation—significantly expands enforcement and prioritizes government-flagged content. This involves medium government pressure combined with moderate platform independence, likely resulting in a constitutional violation. While Twitter has some editorial interest, government involvement substantially shaped the scope and intensity of enforcement.<sup>58</sup>

The *Murthy* facts likely constitute a constitutional violation under this framework. The government applied sustained, high-level pressure through repeated contacts, expressions of frustration, implicit regulatory threats, and positive reinforcement for compliance.<sup>59</sup> While platforms had some preexisting interest in content moderation, the government's involvement substantially influenced what content was removed, how quickly, and how broadly policies were enforced. The combination of high pressure and only moderate platform independence crosses the constitutional threshold.

Consider first the government's response to foreign interference in elections. Intelligence agencies routinely identify foreign disinformation operations and share this information with platforms. When the Department of Homeland Security alerts Twitter to a coordinated bot network spreading false information about polling locations, Twitter investigates using its existing detection tools and removes accounts that violate its policies against platform manipulation. Under the sliding scale framework, this interaction raises no constitutional concern. The government pressure is minimal—a single information-sharing communication about a genuine security threat from officials whose legitimate function includes protecting election integrity. Platform independence is strong because Twitter already prohibits coordinated inauthentic behavior for business reasons—authentic user engagement benefits the platform—has established processes for identifying such networks, and applies these policies regardless of the content being spread. The government provided information; the platform made an independent judgment to act on it.

Now consider a different scenario involving the same basic fact pattern but with sustained pressure. Suppose the FBI not only shares information about foreign bot networks but also begins flagging individual posts from authentic American users expressing skepticism about election security. FBI officials hold weekly meetings with platform representatives, provide lists of accounts to review, express frustration when platforms don't act on all flagged content, and reference the platforms' obligation to protect election integrity. Platforms, concerned about maintaining a cooperative relationship with law enforcement and aware of potential regulatory scrutiny, begin removing flagged content even when it doesn't clearly violate existing policies.<sup>60</sup>

Here the framework identifies a constitutional problem. Government pressure has increased substantially—repeated high-level contacts, expectations of compliance, and implicit appeals to

---

<sup>58</sup> Id.

<sup>59</sup> Id.

<sup>60</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024) [Alito, J., dissenting].

regulatory authority. Platform independence has weakened because enforcement extends beyond preexisting policies and responds primarily to government flagging rather than independent identification of problematic content. The government has moved from information sharing to behavior modification, and the platforms' decisions no longer reflect genuine editorial independence.<sup>61</sup>

The framework also addresses situations involving public health emergencies, where government interests are compelling but the risk of censorship is acute. During a pandemic, public health officials have legitimate reasons to combat medical misinformation that could cost lives. When the Surgeon General issues a public advisory about health misinformation and calls on platforms to take action, this constitutes government speech that might influence platform behavior.<sup>62</sup> If platforms respond by modestly adjusting existing health misinformation policies—perhaps by adding COVID-specific examples to existing categories or accelerating review of flagged content—the constitutional calculus depends on the degree of pressure and platform independence. A single public advisory without follow-up pressure, combined with platforms making adjustments clearly within their existing policy framework and business interests, likely does not constitute a violation.

But suppose the scenario escalates. Senior White House officials begin holding regular meetings with platforms, providing specific lists of accounts and posts to remove, praising platforms that comply quickly while publicly criticizing those that don't, and making references to potential changes in Section 230 protections.<sup>63</sup> Platform enforcement dramatically expands, removing not just demonstrably false health claims but also controversial scientific opinions, criticism of public health measures, and reporting on vaccine side effects. Internal platform communications reveal that content is being removed primarily to satisfy government demands rather than because it clearly violates independent editorial standards. Under these facts, the framework identifies a constitutional violation. The pressure is high and sustained, involving senior officials with regulatory authority making implicit threats. Platform independence is weak because enforcement extends far beyond preexisting policies, responds primarily to government pressure rather than independent judgment, and targets content the platforms previously tolerated. Even if the government's public health goals are compelling, the means employed—sustained pressure to censor protected speech—violate the First Amendment.<sup>64</sup>

The framework also clarifies that not all platform responses to government communication implicate constitutional concerns. When government officials speak publicly about issues without directly pressuring platforms, and platforms respond by exercising their editorial judgment, no violation occurs even if the government's speech influenced the platforms' thinking. If the Surgeon General gives speeches about health misinformation and platforms, persuaded by the public health arguments, decide to strengthen their policies, this represents the marketplace of

---

<sup>61</sup> *Id.*

<sup>62</sup> *Moody v. NetChoice, LLC*, 144 S. Ct. at 2383; *NetChoice, LLC v. Paxton*, 144 S. Ct. at 2383.

<sup>63</sup> *NetChoice*, 144 S. Ct. at 2395.

<sup>64</sup> *Miami Herald Publ'g Co. v. Tornillo*, 418 U.S. at 258.

ideas functioning properly. The government used its bully pulpit to advocate for a position, and private actors were persuaded. The absence of direct pressure, threats, or expectations of compliance preserves platform independence. The constitutional line is crossed not when government speech influences private decisions, but when government pressure compromises the genuinely private nature of those decisions.<sup>65</sup>

The framework's implications extend beyond social media to other contexts where the government seeks to influence private actors' speech decisions. Consider university administrators pressuring student newspapers, elected officials threatening to withdraw funding from museums that display controversial art, or police departments requesting that protest organizers change their messaging. In each context, the sliding scale applies. The degree of government pressure—going from suggestion to threat—must be weighed against the private actor's independent interest in the speech decision. When a university president casually suggests that a student newspaper consider the campus climate when making editorial decisions, minimal pressure combined with strong editorial independence raises no constitutional concern. When that same president threatens to withdraw funding unless the newspaper kills a story, high pressure combined with compromised independence—the threat creates a financial incentive contrary to editorial judgment—constitutes state action.<sup>66</sup>

The framework also addresses the temporal dimension of government influence. Government pressure that has ceased may nonetheless create lasting effects on platform behavior. If the White House engaged in sustained pressure that caused platforms to adopt new policies, those policies may remain in effect long after the pressure ends. Platforms may continue enforcing government-preferred content moderation even when no longer under active pressure, either because the policies became institutionalized or because platforms fear renewed pressure if they relax enforcement. The framework recognizes this problem by examining not just current government conduct but also the historical relationship between government pressure and platform policy development. When a platform's current content moderation practices can be traced to earlier government pressure rather than independent editorial judgment, the constitutional violation persists even if active pressure has ceased. This prevents the government from achieving through temporary pressure what the First Amendment would forbid through permanent regulation.<sup>67</sup>

One important limitation of the framework deserves emphasis. It does not require platforms to host any particular speech or prohibit them from moderating content based on their own judgment. Platforms retain full editorial discretion to establish and enforce content policies reflecting their values and business interests. The framework intervenes only when government pressure substantially influences those decisions. A platform that independently chooses to prohibit certain categories of speech remains free to do so, even if users or government officials

---

<sup>65</sup> Id.

<sup>66</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>67</sup> Id.

disagree with that choice. The First Amendment protects platforms' editorial autonomy as much as it protects users' speech. The framework seeks to preserve both by identifying when government involvement compromises the independence that makes platform moderation a protected editorial choice rather than state action.<sup>68</sup>

## VI. Conclusion

The First Amendment does not exist in a binary world, yet state action doctrine has long treated it as if it does. Government conduct is either coercion or persuasion, state action or private choice, constitutional violation or permissible activity. This binary framework served adequately when government-private actor relationships were simpler and more episodic. But the modern relationship between government officials and social media platforms—characterized by sustained engagement, implicit pressure, and intertwined interests—defies such categorical treatment.<sup>69</sup>

*Murthy v. Missouri* exposed the inadequacy of applying binary analysis to spectrum problems. The government's conduct was neither pure persuasion nor explicit coercion.<sup>70</sup> It was sustained, high-level pressure operating through repeated contacts, expressions of official displeasure, veiled regulatory threats, and praise for compliance. Platforms' responses reflected neither pure editorial independence nor complete capitulation. They enforced existing policies more aggressively, expanded enforcement in government-prioritized areas, and responded with particular alacrity to government-flagged content. The binary framework had no vocabulary for analyzing this middle ground, leaving the Court unable to reach the constitutional merits and forcing dismissal on standing grounds.<sup>71</sup>

The sliding scale framework proposed in this Article provides that missing vocabulary. By assessing the degree of government pressure against the strength of platform independence, courts can identify when government involvement crosses the line from legitimate communication to unconstitutional censorship. The framework recognizes that both factors operate on continuums and that constitutional violations emerge not from any single element but from the combination of substantial pressure and compromised independence. This approach respects all three constitutional interests at stake: the government's ability to communicate about matters of public concern, platforms' editorial autonomy, and users' freedom from government-compelled censorship.<sup>72</sup>

This framework is not without limitations. It requires factual inquiry into both government conduct and platform motivations, demanding more of courts than a simple binary determination. It involves judgment calls about where conduct falls on the pressure spectrum and how

---

<sup>68</sup> Id.

<sup>69</sup> U.S. CONST. amend. I.

<sup>70</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>71</sup> Id.

<sup>72</sup> Communications Decency Act of 1996, 47 U.S.C. § 230 (2018).

independently platforms would have acted. These challenges are real but not insurmountable. Courts regularly make such contextual judgments in other First Amendment contexts, weighing multiple factors to determine constitutional outcomes. The alternative—clinging to a binary framework that leaves vast swathes of government influence beyond constitutional scrutiny—is worse.

As governments and platforms continue to grapple with online misinformation, foreign interference, and harmful content, the constitutional stakes will only grow. Officials will continue seeking to influence platform moderation, and platforms will continue making decisions that affect public discourse. Without a framework capable of identifying when influence becomes coercion, the First Amendment's protections will erode. The government will learn that sustained pressure short of explicit threats achieves censorship without constitutional consequence.<sup>73</sup> Platforms will learn that cooperating with government preferences insulates them from regulatory risk. And users will find their speech suppressed not through formal government action but through a partnership between public pressure and private power that current doctrine cannot reach.<sup>74</sup>

The sliding scale framework offers a path forward. It preserves space for legitimate government-platform cooperation while protecting against government overreach. It respects platforms' editorial rights while ensuring those rights cannot be weaponized to circumvent the First Amendment. Most importantly, it protects users' speech by recognizing that constitutional violations can occur even when government pressure operates subtly and platforms retain nominal decision-making authority. In an age when digital platforms mediate much of public discourse and the government seeks to shape that discourse, the First Amendment demands no less.<sup>75</sup>

---

<sup>73</sup> *Murthy v. Missouri*, 603 U.S. \_\_\_, 144 S. Ct. 2383 (2024).

<sup>74</sup> *Id.*

<sup>75</sup> *Id.*