

MAKE AMERICA FAKE AGAIN?: BANNING DEEPFAKES OF FEDERAL CANDIDATES IN POLITICAL ADVERTISEMENTS UNDER THE FIRST AMENDMENT

*Sophie Loewenstein**

In recent years, artificial intelligence (AI) technology has developed rapidly. Accompanying this advancement in sophistication and accessibility are various societal benefits and risks. For example, political campaigns and political action committees have begun to use AI in advertisements to generate deepfakes of opposing candidates to influence voters. Deepfakes of political candidates interfere with voters' ability to discern falsity from reality and make informed decisions at the ballot box. As a result, these deepfakes pose a threat to the integrity of elections and the existence of democracy. Despite the dangers of deepfakes, regulating false political speech raises significant First Amendment questions.

This Note considers whether the Protect Elections from Deceptive AI Act, a proposed federal ban of AI-generated deepfakes portraying federal candidates in political advertisements, is constitutional. This Note concludes that the bill is constitutional under the First Amendment and that less speech restrictive alternatives fail to address the risks of deepfakes. Finally, this Note suggests revisions to narrow the bill's application and ensure its apolitical enforcement.

| | |
|------------------------------------------------------------------|-----|
| INTRODUCTION | 275 |
| I. DEEPFAKES AND THE FIRST AMENDMENT: A FEDERAL SOLUTION..... | 277 |
| A. <i>What Are “Deepfakes”?</i> | 277 |
| B. <i>Federal Efforts to Regulate Political Deepfakes.</i> | 279 |
| 1. Proposed Federal Legislation | 280 |
| 2. Proposed Agency Action | 282 |

* J.D. Candidate, 2025, Fordham University School of Law; B.A., 2019, Tufts University. Thank you to Professor Abner Greene for his insight and guidance throughout this process. I would also like to thank my primary editor, Jeffrey Wu, and the other members of the *Fordham Law Review* for their assistance and feedback. Lastly, thank you to my friends and family, especially my parents, Myra and David, for their constant love, encouragement, and support.

| | |
|---------------------------------------------------------------------------------------------------------------------------|-----|
| C. <i>The First Amendment: Background, Politics, and Falsity</i> | 283 |
| 1. Constitutional Background..... | 283 |
| 2. First Amendment Rationales | 285 |
| a. <i>Self-Governance</i> | 285 |
| b. <i>Marketplace of Ideas: Counterspeech</i> | 286 |
| c. <i>Negative Theory: Distrust in Government</i> | 286 |
| 3. First Amendment Protections: Political Speech and False Speech | 287 |
| a. <i>Political Speech</i> | 287 |
| b. <i>Falsity Alone and Alvarez</i> | 289 |
| 4. Categories of Unprotected False Speech | 290 |
| a. <i>Defamation</i> | 291 |
| b. <i>Fraud</i> | 292 |
| c. <i>Impersonation of Government Officials</i> | 292 |
| II. IS A FEDERAL BAN OF POLITICAL ADVERTISEMENT DEEPFAKES OF FEDERAL CANDIDATES COMPATIBLE WITH THE FIRST AMENDMENT?..... | 293 |
| A. <i>The Proposed Ban Is Unconstitutional Under the First Amendment</i> | 293 |
| 1. <i>Falsity Alone: The Ban Under Alvarez</i> | 294 |
| 2. <i>Unconstitutional Limitation of Political Speech</i> | 298 |
| 3. <i>The Proposed Ban Is Not Narrowly Tailored</i> | 300 |
| 4. <i>Distrust in Government to Enforce the Ban Apolitically</i> | 302 |
| B. <i>The Proposed Ban Is Constitutional Under the First Amendment</i> | 303 |
| 1. <i>More Than “Falsity Alone”: Fitting the Ban into the Alvarez Framework</i> | 304 |
| 2. <i>Justifying Limiting Political Speech</i> | 308 |
| 3. <i>The Proposed Ban Is Narrowly Tailored</i> | 310 |
| 4. <i>Safeguards Mitigate Partisan Enforcement Concerns</i> | 312 |
| III. IS IT CONSTITUTIONAL?: A NARROW SOLUTION TO A LEGALLY COGNIZABLE HARM..... | 312 |
| A. <i>The Proposed Ban Is Constitutional</i> | 312 |
| B. <i>Less Speech Restrictive Alternatives Are Ineffective</i> | 315 |
| C. <i>Recommended Revisions to the Protect Elections from Deceptive AI Act</i> | 317 |
| CONCLUSION..... | 320 |

INTRODUCTION

In May 2023, the Republican National Committee created an advertisement using artificial intelligence¹ (AI)-generated deepfake² imagery depicting stark scenes of a hypothetical second term for President Biden.³ One frame of the video realistically portrayed a fictionalized scene of a locked-down San Francisco that was patrolled by the military due to the fentanyl crisis and crime.⁴ The top left corner of the advertisement stated, “Built entirely with AI imagery.”⁵ In June 2023, Republican presidential candidate Ron DeSantis’s campaign ran an advertisement that showed AI-generated images of former President Donald J. Trump hugging and kissing Dr. Anthony S. Fauci.⁶ The fake pictures were interposed with real photos and audio of President Trump and Dr. Fauci, further blurring the lines between falsity and reality.⁷

Swift advancements in AI technology have allowed users to generate increasingly realistic audio, images, and videos.⁸ As a result of these improvements, people are struggling to differentiate between real and AI-generated media.⁹ To further complicate the issue, studies have shown that people not only fail to distinguish between genuine and fabricated media, but also view AI-synthesized faces as more trustworthy.¹⁰ Now, politicians are using deepfakes in campaign advertisements to make their opponents appear to do or say things that never happened.¹¹ As the public gains

1. “Artificial intelligence is a field of science concerned with building computers and machines that can reason, learn, and act in such a way that would normally require human intelligence” *What Is Artificial Intelligence (AI)?*, GOOGLE CLOUD, <https://cloud.google.com/learn/what-is-artificial-intelligence> [<https://perma.cc/6UA9-4WKZ>] (last visited Aug. 31, 2024).

2. A deepfake is “an image or recording that has been convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said.” *Deepfake*, MERRIAM-WEBSTER, <https://www.merriamwebster.com/dictionary/deepfake> [<https://perma.cc/A6KM-GYEG>] (last visited Aug. 31, 2024).

3. GOP, *Beat Biden*, YOUTUBE (Apr. 25, 2023), <https://www.youtube.com/watch?v=kLMMxgtxQ1Y> [<https://perma.cc/6VDB-C7SU>].

4. *Id.*

5. *Id.*

6. @DeSantisWarRoom, X (June 5, 2023, 3:13 PM), <https://x.com/DeSantisWarRoom/status/1665799058303188992> [<https://perma.cc/Z7MD-F7LH>].

7. *See id.* However, context is added below the video posted on X, which states, “[I]n this video, a collage of photos of President Trump hugging Anthony Fauci appear. These pictures are not real; they are AI-generated images.” *Id.*

8. TODD C. HELMUS, RAND CORP., ARTIFICIAL INTELLIGENCE, DEEPFAKES, AND DISINFORMATION 1–2 (2022), https://www.rand.org/content/dam/rand/pubs/perspectives/PEA1000/PEA10431/RAND_PEA1043-1.pdf [<https://perma.cc/3MFV-KHYF>].

9. *See id.* at 3.

10. SOPHIE J. NIGHTINGALE & HANY FARID, PNAS, AI-SYNTHESIZED FACES ARE INDISTINGUISHABLE FROM REAL FACES AND MORE TRUSTWORTHY 2 (2022), <https://www.pnas.org/doi/pdf/10.1073/pnas.2120481119> [<https://perma.cc/2A7P-YZ7L>].

11. *See, e.g., Beat Biden*, *supra* note 3; Tatyana Monnay, *Deepfake Political Ads Are “Wild West” for Campaign Lawyers*, BLOOMBERG L. (Sept. 5, 2023, 5:00 AM), <https://news.bloomberglaw.com/business-and-practice/deepfake-political-ads-are-wild-west-for-camp>

familiarity with deepfakes, candidates caught on video can exploit this confusion and claim that a real video is a deepfake.¹² Notably, these consequences are exacerbated by some political campaigns using political microtargeting techniques (PMTs).¹³ PMTs gather information about voters and target them with advertisements tailored to their preferences.¹⁴ As a result, campaigns can use deepfake advertisements to target subgroups of voters that are more susceptible to believing that the fraudulent videos, images, or audio are real.¹⁵ All populations exhibit vulnerability to deepfakes, but this vulnerability varies with “age, political orientation, and trust in information sources.”¹⁶

Although the dangers of this technology have become more apparent, attempts to regulate it have struggled to come to fruition. As of July 2024, forty-four states and the District of Columbia have proposed, passed, or declined to pass legislation that limits the use of political deepfakes.¹⁷ Scholars and legislators have considered other potential avenues of regulation, including defamation, privacy, and copyright law.¹⁸ Proposed federal legislation, however, has failed to gain traction.¹⁹ Though Congress has declined to advance legislation, parties have submitted petitions requesting the Federal Election Commission (FEC) to consider implementing regulations of political advertisements that use deepfakes.²⁰ Absent limitation, unfettered use of deepfakes will be used to influence voters and undermine the integrity of elections.²¹ However, there are First Amendment

aign-lawyers [<https://perma.cc/H3FF-YZXH>] (“A clip circulating on social media of Sen. Elizabeth Warren (D-Mass.) manipulated Warren as saying GOP votes ‘could threaten the integrity of the election.’”).

12. See Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1758 (2019) (“[D]eep fakes make it easier for liars to avoid accountability for things that are in fact true.”).

13. See Tom Dobber, Nadia Metoui, Damian Trilling, Natali Helberger & Claes De Vreese, *Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?*, 26 INT’L J. OF PRESS/POL. 69, 70 (2021).

14. *Id.*

15. *Id.* at 72.

16. See CHRISTOPHER DOSS, JARED MONDSCHNEIN, DULE SHU, TAL WOLFSON, DENISE KOPECKY, VALERIE A. FITTON-KANE, LANCE BUSH & CONRAD TUCKER, *SCI. REPS., DEEPFAKES AND SCIENTIFIC KNOWLEDGE DISSEMINATION 2* (2023), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10439167/pdf/41598_2023_Article_39944.pdf [<https://perma.cc/82C2-5PML>].

17. See *Tracker: State Legislation on Deepfakes in Elections*, PUB. CITIZEN, <https://www.citizen.org/article/tracker-legislation-on-deepfakes-in-elections/> [<https://perma.cc/74SS-YMFN>] (last visited Aug. 31, 2024).

18. See David Greene, *We Don’t Need New Laws for Faked Videos, We Already Have Them*, ELEC. FRONTIER FOUND. (Feb. 13, 2018), <https://www.eff.org/deeplinks/2018/02/we-dont-need-new-laws-faked-videos-we-already-have-them> [<https://perma.cc/PW33-ZVW8>].

19. See *infra* Part I.B.1.

20. See *infra* Part I.B.2.

21. See Rebecca Green, *Counterfeit Campaign Speech*, 70 HASTINGS L.J. 1445, 1457–67 (2019).

questions regarding the constitutionality of regulating deepfake political advertisements.²²

This Note explores whether a federal prohibition of AI-generated deepfakes of federal candidates in political advertisements is constitutional. Part I of this Note explains what deepfakes are and provides a background of the First Amendment. Part II evaluates First Amendment arguments in favor of and against a federal ban of deepfake political advertisements based on the proposed Protect Elections from Deceptive AI Act.²³ Part III concludes that the proposed ban is constitutional, explains why less speech restrictive solutions are inadequate, and recommends changes to the proposed bill.

I. DEEPFAKES AND THE FIRST AMENDMENT: A FEDERAL SOLUTION

This part provides a background on deepfakes and the First Amendment. Part I.A explains deepfake technology and why deepfakes are harmful. Part I.B reviews proposed federal legislation and agency action that aim to regulate deepfakes in elections. Part I.C examines the foundational theories underlying the First Amendment and categories of speech that are constitutionally protected and unprotected.

A. What Are “Deepfakes”?

Deepfakes are digitally created audio or visual media that appear to be authentic but are not.²⁴ The term “deepfake” stems from merging the terms “deep learning” and “fake.”²⁵ Deep learning is a type of machine learning that uses algorithms called artificial neural networks (ANNs).²⁶ ANNs mimic the human brain’s structure and function and can learn from large unstructured audiovisual data sets found online.²⁷ Content generated by deep learning is high quality and appears to be authentic.²⁸ The most realistic deepfakes “require thousands of images to train algorithms to recognize and then manipulate a face.”²⁹ Other audiovisual manipulations are distinguishable from deepfakes, often called “cheap fakes,” or “shallow fakes.”³⁰ Cheap fakes and shallow fakes are created using more inexpensive

22. See Cass R. Sunstein, *Falsehoods and the First Amendment*, 33 HARV. J.L. & TECH. 387, 418–21 (2020).

23. S. 2770, 118th Cong. (2024).

24. RAINA DAVIS, CHRIS WIGGINS & JOAN DONOVAN, BELFER CTR. FOR SCI. AND INT’L AFFS., TECH FACTSHEETS FOR POLICYMAKERS: DEEPFAKES I (2020), <https://www.belfercenter.org/sites/default/files/2020-10/tappfactsheets/Deepfakes.pdf> [<https://perma.cc/W953-PEK3>].

25. *Id.* at 2.

26. *Id.* at 3.

27. *Id.*

28. *Id.* at 2.

29. Soojin Jeong, Margaret Sturtevant & Karis Stephen, *Responding to Deepfakes and Disinformation*, REGUL. REV. (Aug. 14, 2021), <https://www.theregreview.org/2021/08/14/saturday-seminar-responding-deepfakes-disinformation/> [<https://perma.cc/VDF6-VCHX>].

30. DAVIS ET AL., *supra* note 24, at 2.

and accessible methods, such as Photoshop, lookalikes, recontextualizing footage, speeding, or slowing.³¹

An image or video deepfake is usually created using a generative adversarial network (GAN), which is comprised of a pair of ANNs.³² A sample is put into one network, called the generator, which then creates an image.³³ The generator's output is put into the second network, called the discriminator, which is trained by data to categorize an image as real or fake.³⁴ As the discriminator reviews the forged image, it adjusts its feedback until it cannot discern between the forged image and the authentic image.³⁵ The more data that is available for the discriminator to review, the more realistic the resulting image will be.³⁶

An audio deepfake creates an artificial voice from a real person's voice using neural networks trained by speech data sets.³⁷ Deepfake voices are produced by "training machine learning models on large amounts of audio data to mimic a specific speaker."³⁸ This requires manipulating existing footage and recordings of a person's voice and face when they speak.³⁹

This Note will focus on regulating AI-generated deepfakes in political advertisements in which candidates' voices or images are intentionally faked to deceive voters. "Cheap fakes" or "edited-but-real" media causes harm to voters as well,⁴⁰ but they are more easily verifiable.⁴¹ In circumstances in which the source material is real and the fake can be debunked, voters are encouraged to get context from the news, campaigns, or media.⁴² Conversely, with AI-generated deepfakes, voters lack the means to distinguish reality from falsity.⁴³ The proliferation of AI technology to create deepfakes has spurred efforts to regulate its use.⁴⁴ In recent years, passed and proposed state, congressional, and federal agency action has attempted to specifically target the use of political deepfakes.⁴⁵

31. *Id.*

32. *Id.* at 3.

33. Karen Hao, *Inside the World of AI that Forges Beautiful Art and Terrifying Deepfakes*, MIT TECH. REV. (Dec. 1, 2018), <https://www.technologyreview.com/2018/12/01/138847/inside-the-world-of-ai-that-forges-beautiful-art-and-terrifying-deepfakes/> [<https://perma.cc/2RP5-F5LA>].

34. *Id.*

35. DAVIS ET AL., *supra* note 24, at 3.

36. *Id.*

37. *Everything You Need to Know About Deepfake Voice*, MURF AI (Apr. 28, 2023), <https://murf.ai/resources/deepfake-voices/> [<https://perma.cc/2RP5-F5LA>].

38. *Id.*

39. *Id.*

40. See Green, *supra* note 21, at 1452–53; see, e.g., Matthew Brown, *Fact Check: Video of Speaker Nancy Pelosi Altered, Selectively Edited*, USA TODAY (Aug. 11, 2020, 5:37 PM), <https://www.usatoday.com/story/news/factcheck/2020/08/11/fact-check-video-pelosi-altered-and-selectively-edited/3332920001/> [<https://perma.cc/DBE4-79HA>].

41. See Green, *supra* note 21, at 1452–53.

42. *Id.*

43. *Id.*

44. See *infra* Part I.B.

45. See *infra* Parts I.B.1–2.

B. Federal Efforts to Regulate Political Deepfakes

This section focuses on proposed congressional and agency actions aimed at regulating the use of deepfakes to influence voters and undermine elections. This Note's scope is limited to federal solutions. However, California,⁴⁶ Michigan,⁴⁷ Minnesota,⁴⁸ Texas,⁴⁹ and Washington⁵⁰ were some of the first states to enact laws seeking to limit deceptive political deepfakes in the context of elections. Subsequently, forty-four states and the District of Columbia have also contemplated similar legislation.⁵¹ Others have suggested that instead of creating new laws, existing laws—such as defamation or copyright laws—should be applied to political deepfakes.⁵²

46. In 2020, California enacted legislation that prohibits the distribution of materially deceptive media of a candidate “with actual malice” within sixty days of an election. CAL. ELEC. CODE § 20010(a) (West 2019). To establish liability, a person must intend “to injure the candidate’s reputation or to deceive a voter into voting for or against the candidate.” *Id.* The statute establishes civil penalties through injunctive relief or damages. *Id.* § 20010(c).

47. In 2023, Michigan passed legislation that criminalizes the dissemination of AI-generated deepfakes of political candidates without a disclosure. MICH. COMP. LAWS ANN. § 168.932f (West 2024). To impose liability, the actor must circulate the deepfake within ninety days of an election. *Id.* § 168.932f(1). The law requires that the actor to intend for the deepfake’s distribution to harm a candidate’s “reputation or electoral prospects,” and that the distribution be “reasonably likely to cause that result.” *Id.* Additionally, the actor must “intend[] the [deepfake’s] distribution to change the voting behavior of electors . . . by deceiving the electors into incorrectly believing that the [candidate] . . . engaged in the speech or conduct depicted, and the distribution is reasonably likely to cause that result.” *Id.*

48. In 2023, Minnesota approved a law that criminalizes the knowing dissemination of a deepfake with the intent “to injure a candidate or influence the result of an election.” MINN. STAT. ANN. § 609.771, subdiv. 2 (West 2023). In 2024, the statute was amended to include that it applied to deepfakes disseminated “within 90 days before a political party nominating convention” or “after the start of the absentee voting period prior to a presidential nomination primary, or a regular or special state or local primary or general election.” 2024 Minn. Sess. Law Serv. 40–42 (West).

49. In 2019, Texas passed legislation that criminalizes creating or distributing a deepfake within thirty days of an election. TEX. ELEC. CODE ANN. § 255.004(d) (West 2023). It requires the creator to act with the “intent to injure a candidate” or “influence the result of an election.” *Id.* However, in May 2023, a Texas Court of Appeals held that the statute was unconstitutional. *See Ex parte Stafford*, 667 S.W.3d 517, 532 (Tex. Ct. App. 2023).

50. In 2023, Washington enacted a law that requires clear disclosure when a person uses deepfakes in election-related media. WASH. REV. CODE § 42.62.020 (2023). This law permits candidates to seek injunctive or other equitable relief prohibiting the publication of the synthetic media. *Id.* § 42.62.020(2). Moreover, it permits individuals to raise an affirmative defense by including a disclosure stating the media was manipulated. *Id.* § 42.62.020(4). In May 2024, the statute was slightly amended. *See* 2024 Wash. Sess. Laws 19, 116.

51. *See Tracker: State Legislation on Deepfakes in Elections*, *supra* note 17.

52. *But see* Anna Pesetski, Note, *Deepfakes: A New Content Category for a Digital Age*, 29 WM. & MARY BILL RTS. J. 503, 525 (2020) (“[E]xisting laws are not enough to stop deepfakes because they do not effectively address the dangers that deepfakes pose.”). Scholars have argued that a federal statute is a superior remedy to utilizing preexisting defamation law because libel suits are next to impossible to win and are cost and time prohibitive, relief likely cannot be granted until after the election, it is difficult to prove personal damages, and proposals to limit political deepfakes are not based on protecting a candidate’s reputation, but on the public’s compelling interest in reliable elections. *See* Jason Zenor, *A Reckless Disregard for Truth?: The Constitutional Right to Lie in Politics*, 38 CAMPBELL L. REV. 41, 69 (2016).

However, for federal elections, a patchwork, state-by-state solution would not only be ineffective and inconsistent, but also challenging to enforce.⁵³

This section focuses on proposed federal regulatory efforts through Congress and the FEC.⁵⁴ Part I.B.1 reviews proposed federal legislation seeking to limit the harms of AI-generated political deepfakes. Part I.B.2 details suggested agency action through the FEC's rulemaking authority.

1. Proposed Federal Legislation

Thus far, Congress has not passed legislation targeting deepfakes in the campaign setting. However, previously introduced legislation has sought to regulate deepfakes outside of politics.⁵⁵ Recently, efforts have begun to specifically target deepfakes in the context of elections and political advertisements.

In 2020, Representative Stephen F. Lynch introduced the Deepfakes in Federal Elections Prohibition Act.⁵⁶ The bill would criminalize the distribution, with "actual malice,"⁵⁷ of materially deceptive AI-generated audio or visual media depicting a federal candidate within sixty days of a federal election.⁵⁸ To be liable, the actor must (1) have knowledge or reckless disregard as to the media's falsity and (2) intend "to injure the candidate's reputation or to deceive a voter into voting for or against the candidate."⁵⁹ The proposed bill also permits a candidate "whose voice or likeness appears" in the deceptive media to bring an action for damages.⁶⁰ However, this bill has not received a hearing in a congressional committee.⁶¹

53. Nicholas O'Donnell, Note, *Have We No Decency?: Section 230 and the Liability of Social Media Companies for Deepfake Videos*, 2021 U. ILL. L. REV. 701, 717 ("Even if all states passed deepfake legislation, the subsequent regime would almost certainly lack comprehensiveness and simplicity, as each would define the problem differently and implement substantially different solutions.").

54. This Note does not consider barriers to online platform liability regarding regulating deepfake political advertisements. For example, § 230 of the Communications Decency Act, 47 U.S.C. § 230 (1996), provides online platforms with a liability shield for hosting harmful content. See Chesney & Citron, *supra* note 12, at 1795–801. Given the accessibility of AI technology, online users can create deepfakes of candidates. See *id.* at 1762–63. However, questions regarding platform liability are beyond the scope of this Note.

55. See, e.g., Malicious Deep Fake Prohibition Act of 2018, S. 3805, 115th Cong. (2018); Deep Fakes Accountability Act, H.R. 3230, 116th Cong. (2019); Deepfake Report Act of 2019, S. 2065, 116th Cong. (2019).

56. H.R. 6088, 116th Cong. (2020).

57. See *N.Y. Times Co. v. Sullivan*, 376 U.S. 279–81 (1964) (defining "actual malice").

58. H.R. 6088, § 2. A violator of the Act is subject to a fine, up to five years in prison, or both. *Id.*

59. *Id.*

60. *Id.*

61. Douglas Mirell, *Deepfakes Remain a Threat Ahead of 2024 Elections*, JD SUPRA (Sept. 19, 2023), <https://www.jdsupra.com/legalnews/deepfakes-remain-a-threat-ahead-of-2024-6429332/> [<https://perma.cc/9RTQ-EF8Y>].

In May 2023, Senators Amy Klobuchar, Cory Booker, and Michael Bennet,⁶² along with Representative Yvette D. Clark,⁶³ presented the Require the Exposure of AI-Led (REAL) Political Advertisements Act.⁶⁴ This would amend the Federal Election Campaign Act of 1971⁶⁵ (FECA) to give the FEC authority to regulate the use of AI in campaign advertisements.⁶⁶ The bill would require a disclosure if AI was used to generate any media in the advertisement.⁶⁷ Some have argued that this bill is likely to fail in the Republican-controlled House of Representatives.⁶⁸

In July 2023, Representative Adriano Espaillat proposed the Candidate Voice Fraud Prohibition Act.⁶⁹ The bill criminalizes the use of AI-generated audio to impersonate a candidate's voice without a disclosure in political communications.⁷⁰ Specifically, it seeks to amend the FECA to prohibit the distribution, with "actual malice," of materially deceptive AI-generated audio impersonating a candidate's voice.⁷¹ It requires the intent to "injure the candidate's reputation or to deceive a voter into voting against the candidate."⁷² The bill's application is limited to within ninety days of a general election or sixty days of a primary election.⁷³

In September 2023, Senators Amy Klobuchar, Josh Hawley, Chris Coons, and Susan Collins,⁷⁴ along with Representatives Derek Kilmer and Tony

62. Press Release, Amy Klobuchar, Sen., U.S. Senate, Klobuchar, Booker, Bennet Introduce Legislation to Regulate AI-Generated Content in Political Ads (May 15, 2023), <https://www.klobuchar.senate.gov/public/index.cfm/2023/5/klobuchar-booker-bennet-introduce-legislation-to-regulate-ai-generated-content-in-political-ads> [https://perma.cc/4UQU-G5MH].

63. Press Release, Yvette D. Clarke, Rep., U.S. House of Representatives, Clarke Introduces Legislation to Regulate AI in Political Advertisements (May 2, 2023), <https://clarke.house.gov/clarke-introduces-legislation-to-regulate-ai-in-political-advertisements/> [https://perma.cc/9A7J-5LRQ].

64. S. 1596, 118th Cong. (2023).

65. Pub. L. No. 92-225, 86 Stat. 3 (codified at 52 U.S.C. §§ 30101–30126, 30141–30145).

66. The legislation notes that the FEC shall promulgate regulations for determining whether an advertisement contains AI-generated media, the contents of the disclosure requirement, and a definition for AI-generated content under the Act. S. 1596, § 4.

67. *Id.*

68. *See* Monnay, *supra* note 11.

69. H.R. 4611, 118th Cong. (2023); Press Release, Adriano Espaillat, Rep., U.S. House of Representatives, Representative Adriano Espaillat Introduces Legislation Regulating the Influence of Artificial Intelligence (AI) on U.S. Elections (July 18, 2023), <https://espaillat.house.gov/media/press-releases/representative-adriano-espaillat-introduces-legislation-regulating-influence> [https://perma.cc/8CHY-7YDW].

70. Criminal penalties include a fine, imprisonment for no more than two years, or both. H.R. 4611, § 3.

71. *Id.* The legislation notes that the FEC shall promulgate regulations to determine whether the materially deceptive audio is satire or parody, and what constitutes a sufficient disclosure. *Id.*

72. *Id.*

73. *Id.*

74. Press Release, Amy Klobuchar, Sen., U.S. Senate, Klobuchar, Hawley, Coons, Collins Introduce Bipartisan Legislation to Ban the Use of Materially Deceptive AI-Generated Content in Elections (Sept. 12, 2023), <https://www.klobuchar.senate.gov/public/index.cfm/2023/9/klobuchar-hawley-coons-collins-introduce-bipartisan-legislation-to-ban-the-use-of-materially-deceptive-ai-generated-content-in-elections> [https://perma.cc/8S8H-WP6Q].

Gonzales,⁷⁵ introduced the Protect Elections from Deceptive AI Act.⁷⁶ This bill seeks to amend the FECA to prohibit the knowing distribution of materially deceptive AI-generated audio, images, or video relating to federal candidates in political advertisements.⁷⁷ In May 2024, the Senate Committee on Rules and Administration advanced the bill.⁷⁸ The bill permits candidates targeted by misleading content to have that content taken down and seek damages in federal court.⁷⁹ The ban extends to any person, political action committee (PAC), or other entity that distributes the deceptive content and intends to “influenc[e] an election” or “solicit[] funds.”⁸⁰

2. Proposed Agency Action

The FEC has issued a notice seeking public comment on commencing a formal rulemaking process on whether fraudulent misrepresentations should apply to deepfakes in campaign advertisements.⁸¹ On May 16, 2023, Public Citizen submitted a petition requesting the FEC to engage in rulemaking to clarify an existing law on fraudulent misrepresentation.⁸² The petition asked the FEC to determine whether legislation targeting fraudulent misrepresentation applies to deliberately deceptive AI-generated campaign advertisements.⁸³ According to FECA and the FEC’s implementing regulations,⁸⁴ political candidates may not “fraudulently misrepresent” themselves as speaking or acting for or on behalf of another candidate or political party on a matter damaging to the other candidate or political party.⁸⁵

75. Press Release, Derek Kilmer, Rep., U.S. House of Representatives, Kilmer Introduces Legislation to Protect Elections from Deceptive Artificial Intelligence in Political Ads (May 14, 2024), <https://kilmer.house.gov/media/press-releases/kilmer-introduces-legislation-to-protect-elections-from-deceptive-artificial-intelligence-in-political-ads> [https://perma.cc/7ZYT-ESZ5].

76. S. 2770, 118th Cong. (2023). A slightly amended version was introduced on May 15, 2024. Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. (2024).

77. S. 2770, § 2.

78. Press Release, Amy Klobuchar, Sen., U.S. Senate, Klobuchar Statement on Rules Committee Passage of Three Bipartisan AI and Elections Bills (May 15, 2024), <https://www.klobuchar.senate.gov/public/index.cfm/2024/5/klobuchar-statement-on-rules-committee-passage-of-three-bipartisan-ai-and-elections-bills> [https://perma.cc/KDU6-ZQHF].

79. See Press Release, *supra* note 74.

80. S. 2770, § 2.

81. See David Garr, *Comments Sought on Amending Regulation to Include Deliberately Deceptive Artificial Intelligence in Campaign Ads*, FED. ELECTION COMM’N (Aug. 16, 2023), <https://www.fec.gov/updates/comments-sought-on-amending-regulation-to-include-deliberately-deceptive-artificial-intelligence-in-campaign-ads/> [https://perma.cc/67YB-WY3U].

82. See Letter from Robert Weissman, President, Pub. Citizen & Craig Holman, Gov’t Affs. Lobbyist, Pub. Citizen, to Lisa J. Stevenson, Gen. Counsel, FEC (May 16, 2023), <https://www.citizen.org/wp-content/uploads/Petition-for-Rulemaking-52-USC-30124.pdf> [https://perma.cc/J23Z-XSQ7].

83. *Id.* at 1.

84. FEC Contribution and Expenditure Limitations and Prohibitions Rules, 11 C.F.R. § 110.16 (2024). Prosecutions under 11 C.F.R. § 110.16 have primarily been in relation to fraudulent solicitation of funds. See Green, *supra* note 21, at 1470 n.137.

85. 11 C.F.R. § 110.16. The FEC initially declined to seek public comment on Public Citizen’s first petition. See Letter from Robert Weissman, President, Pub. Citizen & Craig

Public Citizen’s second petition requested that the FEC revise 11 C.F.R. § 110.16(a) to state that a candidate’s use of deliberately false AI-generated content in campaign advertisements triggers the fraudulent misrepresentation law.⁸⁶ The petition distinguished instances when fraudulent misrepresentation would and would not apply.⁸⁷ For example, the prohibition would not cover the use of AI generally, but only deepfakes or other similar communications.⁸⁸ It also proposed that the law would not apply when there is a sufficiently prominent disclosure in the advertisement that the media is AI-generated and portrays fictitious statements and actions.⁸⁹ In August 2023, the FEC issued a notice seeking public comment regarding Public Citizen’s petition.⁹⁰

*C. The First Amendment:
Background, Politics, and Falsity*

The harms that AI-generated deepfakes present to democracy and voters have prompted the aforementioned state and federal regulatory efforts. However, questions remain as to whether these proposed actions are constitutional under the First Amendment.⁹¹ Part I.C.1 explains the background of the First Amendment. Part I.C.2 describes the purpose and theoretical justifications for freedom of speech. Part I.C.3 details the protections provided to political and false speech. Part I.C.4 expounds upon categories of unprotected false speech that the U.S. Supreme Court has held may be constitutionally regulated.

1. Constitutional Background

The First Amendment establishes that “Congress shall make no law . . . abridging the freedom of speech.”⁹² The Supreme Court has applied the First Amendment to the states through the Fourteenth Amendment.⁹³ Therefore, the government has “no power to restrict expression because of

Holman, Gov’t Affs. Lobbyist, Pub. Citizen, to Lisa J. Stevenson, Gen. Counsel, FEC 1 (July 13, 2023), <https://www.citizen.org/wp-content/uploads/2nd-Submission-FEC-petition-30124-final-2.pdf> [<https://perma.cc/PJK2-CGNY>]. First, the FEC doubted it had the statutory authority to regulate AI in campaign advertisements. *Id.* Second, Public Citizen did not cite the specific regulation it sought to amend. *Id.* However, the FEC later changed its position following a review of Public Citizen’s amended petition. *See* Garr, *supra* note 81.

86. *See* Letter from Robert Weissman & Craig Holman to Lisa Stevenson, *supra* note 85, at 5.

87. *See id.* at 3–4.

88. *Id.*

89. *Id.*

90. *See* Garr, *supra* note 81.

91. *See* Daniel I. Weiner & Lawrence Norden, *Regulating AI Deepfakes and Synthetic Media in the Political Arena*, BRENNAN CTR. FOR JUST. (Dec. 5, 2023), <https://www.brennancenter.org/our-work/research-reports/regulating-ai-deepfakes-and-synthetic-media-political-arena> [<https://perma.cc/22NK-QD92>].

92. U.S. CONST. amend. I.

93. *Gitlow v. New York*, 268 U.S. 652, 666 (1925).

its message, its ideas, its subject matter, or its content.”⁹⁴ In First Amendment cases, the Supreme Court applies differing levels of scrutiny depending on whether a regulation is content based or content neutral.⁹⁵

A law is facially content based if its text applies to speech based on its subject matter or the viewpoint of the speaker.⁹⁶ A regulation targets subject matter when it limits expression based only on the topic of the speech.⁹⁷ A viewpoint-based regulation targets speech based on the opinion of the speaker on a subject.⁹⁸ The Supreme Court typically views content-based laws as presumptively unconstitutional because they have the potential to exclude specific ideas or speakers from public debate.⁹⁹ As a result, when a law aims to regulate speech based on its content, it typically triggers strict scrutiny.¹⁰⁰

Strict scrutiny requires the government to demonstrate that the law is “necessary” to promote a “compelling government purpose.”¹⁰¹ To withstand strict scrutiny, the legislation may not be too overinclusive or underinclusive.¹⁰² When a law targeting speech is overinclusive, it sweeps too broadly and prohibits protected and unprotected speech.¹⁰³ Conversely, a law that is underinclusive targets some speech for adverse treatment but leaves out other speech that is “indistinguishable in terms of the law’s purpose.”¹⁰⁴ Additionally, if there is a less-restrictive alternative that still serves the government’s purpose, the legislature must use that alternative.¹⁰⁵ Although content-based laws generally result in strict scrutiny, there are speech categories that are subject to lesser scrutiny or may be regulated by the government if certain tests are satisfied.¹⁰⁶

94. *Police Dep’t of Chi. v. Mosley*, 408 U.S. 92, 95 (1972).

95. ERWIN CHERMERINSKY, *CONSTITUTIONAL LAW: PRINCIPLES AND POLICIES* 815 (5th ed., 2015).

96. David L. Hudson Jr., *Content Based*, FREE SPEECH CTR. (Aug. 10, 2023), <https://firstamendment.mtsu.edu/article/content-based> [<https://perma.cc/T7N2-KTMUJ>].

97. *Id.*

98. Kristi Nickodem & Kristina Wilson, *Responding to First Amendment Audits: Content-Based vs. Viewpoint-Based Restrictions*, COATES’ CANONS NC LOC. GOV’T L. (Nov. 18, 2022), <https://canons.sog.unc.edu/2022/11/responding-to-first-amendment-audits-content-based-vs-viewpoint-based-restrictions/> [<https://perma.cc/CW2F-JU8C>].

99. *Reed v. Town of Gilbert*, 576 U.S. 155, 163 (2015).

100. *United States v. Playboy Ent. Grp.*, 529 U.S. 803, 813–14 (2000).

101. CHERMERINSKY, *supra* note 95, at 468.

102. *Id.* at 591.

103. Richard Parker, *Overbreadth*, FREE SPEECH CTR. (Aug. 11, 2023), <https://firstamendment.mtsu.edu/article/overbreadth> [<https://perma.cc/THE6-V5K3>].

104. William E. Lee, *The First Amendment Doctrine of Underbreadth*, 71 WASH. U. L.Q. 637, 637 (1993).

105. *Playboy Ent. Grp.*, 529 U.S. at 813.

106. *See, e.g., N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279–80 (1964) (holding that content-based regulations targeting defamatory statements about public officials made with knowledge that it was false or with reckless disregard of whether it was false are constitutional); *Brandenburg v. Ohio*, 395 U.S. 444, 447–48 (1969) (holding that content-based regulations restricting speech aiming to “incit[e] or produc[e] imminent lawless action” that is “likely to incite or produce such action” are constitutional); *Illinois ex rel. Madigan v. Telemarketing Assocs., Inc.*, 538 U.S. 600, 612 (2003) (noting that the government is permitted to regulate fraudulent speech to prevent public or consumer deception).

In contrast, regulations unrelated to the content of speech—or content-neutral regulations—are subject to an intermediate level of scrutiny.¹⁰⁷ A law is content neutral if it is both viewpoint neutral and subject matter neutral, or if it applies to all speech regardless of the message.¹⁰⁸ Intermediate scrutiny is satisfied if the regulation furthers an “important” or “substantial” government interest and the restriction is narrowly tailored to serve that interest.¹⁰⁹

2. First Amendment Rationales

Three common rationales for the First Amendment include self-governance, the marketplace of ideas, and distrust in the government to regulate. Self-governance and the marketplace of ideas are examples of positive theories of the First Amendment.¹¹⁰ Positive, or affirmative, theory posits that speech is protected because of its affirmative value to democracy or the individual.¹¹¹ On the other hand, negative theory is premised on the risk of limiting speech, such as fear of government censorship.¹¹²

a. Self-Governance

Philosopher Alexander Meiklejohn asserted that the basis of freedom of speech “springs from the necessities . . . of self-government.”¹¹³ To successfully engage in democracy, voters must be knowledgeable regarding government issues.¹¹⁴ Freedom of speech facilitates the democratic process by providing voters with unrestricted access to information regarding candidates, political issues, and social issues so they may make informed decisions.¹¹⁵ Through encouraging this exchange of ideas, it assures that the people may bring about the change that they desire.¹¹⁶

Moreover, the First Amendment serves as a check on the government and permits voters to “retain a veto power” against public officials.¹¹⁷ It encourages open debate of issues, which includes permitting the people and the press to discuss officials and public affairs with impunity.¹¹⁸ Thus, the First Amendment ensures that the public may engage in political discourse

107. CHEMERINSKY, *supra* note 95, at 815, 817.

108. *Id.* at 816.

109. *Turner Broad. Sys. v. Fed. Comm’n Comm’n*, 512 U.S. 622, 662 (1994).

110. See Helen Norton, *Distrust, Negative First Amendment Theory, and the Regulation of Lies*, KNIGHT FIRST AMEND. INST. AT COLUM. UNIV. (Oct. 19, 2022), <https://knightrcolumbia.org/content/distrust-negative-first-amendment-theory-and-the-regulation-of-lies> [https://perma.cc/M9L7-46YG].

111. *See id.*

112. *See id.*

113. ALEXANDER MEIKLEJOHN, *FREE SPEECH AND ITS RELATION TO SELF-GOVERNMENT* 26 (1948).

114. *Id.* at 25.

115. *See* CHEMERINSKY, *supra* note 95, at 811.

116. *See id.*

117. *See id.*

118. *See* *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 296 (1964) (Black, J., concurring).

and criticize the government and its officials, which is the “essence of self-government.”¹¹⁹

b. Marketplace of Ideas: Counterspeech

In *Abrams v. United States*,¹²⁰ Justice Oliver Wendell Holmes, Jr. first referenced the marketplace of ideas theory behind the First Amendment.¹²¹ Justice Holmes’s dissent stated, in part, “[T]he ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market”¹²² The theory derives from the economic marketplace, in which through competition, superior products sell better than lesser products.¹²³

Ultimately, the marketplace of ideas refers to the belief that the acceptance of an idea depends on its competition with other ideas in the marketplace.¹²⁴ This theory asserts that the merit of ideas is based on this competition and not determined by the government or another authority.¹²⁵ When the government limits speech, it interferes with the marketplace of ideas because it places the regulated speech at a disadvantage.¹²⁶ Consequently, it interrupts the public’s ability to freely engage with ideas and shape society.¹²⁷ Therefore, this rationale suggests that the foundation of the First Amendment is that it creates an open exchange of ideas that compete for public acceptance.¹²⁸

c. Negative Theory: Distrust in Government

The First Amendment’s protections are also justified based on negative theory. Professor Frederick Schauer argues that the First Amendment is based on both positive and negative values.¹²⁹ Positive values include public engagement in democracy and debate.¹³⁰ However, over time, Schauer claims that the theory behind the First Amendment has developed more negatively.¹³¹ Based on negative theory, the U.S. Constitution shields speech, not because it is valuable, but because of the risk of the government

119. See *Garrison v. Louisiana*, 379 U.S. 64, 74–75 (1964).

120. 250 U.S. 616 (1919).

121. *Id.* at 630 (Holmes, J., dissenting).

122. *Id.*

123. David Schultz, *Marketplace of Ideas*, FREE SPEECH CTR. (Jan. 1, 2009), <https://firstamendment.mtsu.edu/article/marketplace-of-ideas/> [https://perma.cc/6WAS-P8ES].

124. *Id.*

125. *Id.*

126. *Reed v. Town of Gilbert*, 576 U.S. 155, 176–77 (2015) (Breyer, J., concurring).

127. *Id.*

128. *N.Y. State Bd. of Elections v. Lopez Torres*, 552 U.S. 196, 208 (2008).

129. Frederick Schauer, *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 HARV. L. REV. 1765, 1790–91 (2004).

130. *Id.*

131. *Id.*

abusing its regulatory power.¹³² Thus, some assert that the rationale of protecting speech under the First Amendment is rooted in fear of government censorship.¹³³

3. First Amendment Protections: Political Speech and False Speech

Limiting deepfakes in political advertisements necessarily involves the regulation of political speech and false speech. Therefore, it is essential to consider when these categories of speech are protected under the First Amendment and when they can be regulated.

a. Political Speech

Political speech is “central to the meaning and purpose of the First Amendment.”¹³⁴ It is highly protected not only due to its value to the speaker, but also its value to listeners in exercising self-government.¹³⁵ Legislation limiting deepfakes in political advertisements focuses on a narrow subset of political speech—campaign speech. In political campaigns, the First Amendment provides its “fullest and most urgent application.”¹³⁶

However, what is considered political speech is difficult to define. Speech is a matter of public concern when it relates to any matter of “political, social, or other concern to the community.”¹³⁷ The Supreme Court has defined core political speech as an “interactive communication concerning political change.”¹³⁸ Given the importance of this speech, it is challenging for a law implicating political speech to overcome strict scrutiny.¹³⁹

Although strict scrutiny is difficult to overcome, not all content-based laws targeting political speech fail strict scrutiny review.¹⁴⁰ When evaluating legislation that limits political speech in the election context, the Court has made a distinction between harm to the electoral process and harm regarding electoral substance.¹⁴¹ The Court has determined legislation targeting the

132. Norton, *supra* note 110; *see also* United States v. Alvarez, 567 U.S. 709, 723 (2012) (plurality opinion) (“Permitting the Government to decree this speech to be a criminal offense . . . has no clear limiting principle.”); Cohen v. California, 403 U.S. 15, 26 (1971) (“Indeed, governments might soon seize upon the censorship of particular words as a convenient guise for banning the expression of unpopular views.”).

133. Norton, *supra* note 110.

134. Citizens United v. FEC, 558 U.S. 310, 329 (2010).

135. *See* Martin H. Redish & Julio Pereyra, *Resolving the First Amendment’s Civil War: Political Fraud & the Democratic Goals of Free Expression*, 62 ARIZ. L. REV. 451, 469–70 (2020).

136. *Citizens United*, 558 U.S. at 339 (quoting *Eu v. S.F. Cnty. Democratic Cent. Comm.*, 489 U.S. 214, 223 (1989)).

137. *Connick v. Myers*, 461 U.S. 138, 146 (1983).

138. *Meyer v. Grant*, 486 U.S. 414, 421–22 (1988).

139. *See* *Burson v. Freeman*, 504 U.S. 191, 211–12 (1992).

140. *See, e.g., id.*; *Holder v. Humanitarian L. Project*, 561 U.S. 1, 39–40 (2010).

141. *See AI and the Future of Our Elections: Hearing Before the S. Comm. on Rules & Admin.*, 118th Cong. 6–8 (2023) (prepared statement of Ari Cohn, Counsel, TechFreedom).

electoral process is typically upheld on constitutional grounds.¹⁴² Conversely, legislation targeting harm to electoral substance, or the content of political conversation, is suspect and raises more significant constitutional questions.¹⁴³

Regulations that limit speech based on harm to the electoral process concern the voting procedure itself.¹⁴⁴ The Court has recognized that the government has a compelling interest in preserving the integrity of elections, such as safeguarding access to the ballot.¹⁴⁵ These efforts include protecting voters from undue influence, confusion, and intimidation.¹⁴⁶ For example, the Court has upheld campaign-free zones around polling places¹⁴⁷ and restrictions on who may appear on the ballot.¹⁴⁸ Therefore, regulations limiting the dissemination of knowingly false disinformation about the voting process have been found constitutional.¹⁴⁹

On the other hand, legislation targeting harm to electoral substance is on less certain footing regarding its constitutionality.¹⁵⁰ Electoral conversation laws seek to regulate speech based on its content.¹⁵¹ There is debate among courts as to whether the government has a compelling interest in protecting the public from false or misleading campaign speech.¹⁵² Some courts view permitting the government to restrict this speech as “patronizing and paternalistic” because it assumes that the public is unable to uncover misleading information.¹⁵³ However, other courts have held that the government does have a compelling interest in an informed electorate that is not tainted by deceptive or false campaign speech.¹⁵⁴ Ultimately, whether political speech can be constitutionally limited is dependent on identifying a compelling interest and whether the statute is sufficiently narrowly tailored to withstand strict scrutiny.¹⁵⁵

142. *See id.* at 6–7.

143. *See id.* at 7–8.

144. *See id.* at 6.

145. *See id.*; *see, e.g.*, *Eu v. S.F. Cnty. Democratic Cent. Comm.*, 489 U.S. 214, 231 (1989) (“A State indisputably has a compelling interest in preserving the integrity of its election process.”); *Buckley v. Valeo*, 424 U.S. 1, 26–29 (1976) (noting a compelling government interest in preserving the integrity of the electoral process, preventing corruption, and sustaining the responsibility of citizens in a democracy).

146. *Burson v. Freeman*, 504 U.S. 191, 199 (1992).

147. *See id.*

148. *See Am. Party of Tex. v. White*, 415 U.S. 767, 794–95 (1974).

149. *See AI and the Future of Our Elections*, *supra* note 141, at 6–7 (prepared statement of Ari Cohn, Counsel, TechFreedom).

150. *See id.* at 7–8.

151. *See id.* at 7.

152. *See id.* at 8.

153. *Id.* (citing *State ex rel. Pub. Disclosure Comm’n v. 119 Vote No! Comm.*, 135 Wash. 2d 618, 631–32 (Wash. 1998) (en banc)).

154. *Id.*; *see, e.g.*, *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 473 (6th Cir. 2016); *281 Care Comm. v. Arneson*, 766 F.3d 774, 786–87 (8th Cir. 2014).

155. *AI and the Future of Our Elections*, *supra* note 141, at 8 (prepared statement of Ari Cohn, Counsel, TechFreedom).

b. Falsity Alone and Alvarez

In *United States v. Alvarez*,¹⁵⁶ the Supreme Court evaluated whether the First Amendment protects verifiably false speech. In *Alvarez*, the Court struck down the Stolen Valor Act of 2005,¹⁵⁷ which criminalized knowingly¹⁵⁸ false representations about having been awarded military decorations or honors.¹⁵⁹

In the plurality opinion, the Court held that “falsity alone may not suffice to bring the speech outside the First Amendment.”¹⁶⁰ In exploring this principle, the opinion distinguished between unprotected lies and protected lies.¹⁶¹ The plurality determined false speech can be limited when the speaker intended to cause a “legally cognizable harm” that the First Amendment traditionally does not protect.¹⁶² In applying strict scrutiny, the plurality concluded that the law was not adequately tailored to the harm to be avoided.¹⁶³ The Court reasoned that the government failed to show how the lie would mislead the public or increase public cynicism of military honors.¹⁶⁴ This decision warned that punishing false speech would deter free debate and other less-restrictive solutions, such as counterspeech¹⁶⁵ could overcome the lie.¹⁶⁶

In Justice Stephen G. Breyer’s concurrence, joined by Justice Kagan, they “rejected the plurality’s ‘strict categorical analysis.’”¹⁶⁷ Instead, they suggested that a law targeting harmful falsehoods may be constitutional if it does not “disproportionately damage First Amendment interests.”¹⁶⁸ In evaluating a law restricting false speech, Justice Breyer stated it was necessary to engage in balancing the “seriousness of the speech-related harm the provision will likely cause, nature and importance of the provision’s countervailing objectives, extent to which the provision will tend to achieve those objectives, and whether there are other, less restrictive ways of doing so.”¹⁶⁹ In applying intermediate scrutiny, Justice Breyer concluded that the

156. 567 U.S. 709 (2012).

157. 18 U.S.C. § 704(b).

158. *Alvarez*, 567 U.S. at 740 n.1 (Alito, J., dissenting) (“Although the Act does not use the term ‘knowing’ or ‘knowingly,’ . . . [t]he Act’s use of the phrase ‘falsely represents,’ moreover, connotes a knowledge requirement.” (citation omitted)).

159. *Id.* at 715–16 (plurality opinion).

160. *Id.* at 719.

161. *See id.* at 717–20.

162. *Id.* at 719.

163. *See id.* at 728–30.

164. *See id.* at 727.

165. Counterspeech is described as a “range of responses, critiques, rebuttals, and disapproval that occurs after disfavored, unpopular, offensive, or hateful speech is uttered.” G.S. Hans, *Changing Counterspeech*, 69 CLEV. ST. L. REV. 749, 751 (2021).

166. *Alvarez*, 567 U.S. at 726–29.

167. Richard L. Hasen, *A Constitutional Right to Lie in Campaigns and Elections?*, 74 MONT. L. REV. 53, 66 (2013) (citing *Alvarez*, 567 U.S. at 730 (Breyer, J., concurring)).

168. Chesney & Citron, *supra* note 12, at 1790–91.

169. *Alvarez*, 567 U.S. at 730 (Breyer, J., concurring).

statute was not narrowly tailored to a specific “subset of lies where specific harm is more likely to occur.”¹⁷⁰

Justice Alito, joined by Justice Thomas and Justice Antonin Scalia, dissented.¹⁷¹ Justice Alito asserted he would have upheld the statute to counteract an “epidemic of false claims about military decorations.”¹⁷² Justice Alito argued that false factual statements should not receive First Amendment protections except when necessary to provide sufficient breathing space for true statements on related subjects.¹⁷³

Although the *Alvarez* decision was fractured, the Justices shared a common apprehension.¹⁷⁴ All expressed skepticism regarding laws that target false speech about issues of public concern, such as religion, history, politics, and literature.¹⁷⁵ In these areas, there is a higher risk of government censorship and abuse of power for political ends.¹⁷⁶ Also, the Justices agreed that when false speech causes “legally cognizable harm” to identifiable victims, liability can be imposed without First Amendment concerns.¹⁷⁷

4. Categories of Unprotected False Speech

Although falsity in and of itself does not bring speech outside of the First Amendment, the Court has recognized specific categories of false speech that are afforded limited protection or are wholly unprotected by the First Amendment.¹⁷⁸ These categories are subject to differing standards of review than the typical strict scrutiny applied to content-based regulations.¹⁷⁹ For example, the Court has upheld laws with content-based restrictions in the context of defamation¹⁸⁰ and fraud.¹⁸¹ Lower courts have also upheld content-based laws targeting impersonation of government officials.¹⁸² The Court has determined that the government is permitted to regulate certain categories of false speech because the speech’s harm outweighs its value.¹⁸³ Professor Marc Jonathan Blitz argues that false speech can be unprotected by

170. *Id.* at 736.

171. *See id.* at 739 (Alito, J., dissenting).

172. *Id.*

173. *See id.* at 750.

174. *See Chesney & Citron, supra* note 12, at 1791.

175. *Id.*

176. Hasen, *supra* note 167, at 69.

177. *See* Marc Jonathan Blitz, *Lies, Line Drawing, and (Deep) Fake News*, 72 OKLA. L. REV. 59, 70 (2018).

178. *See* CHEMERINSKY, *supra* note 95, at 853.

179. *See id.*

180. *See, e.g.,* N.Y. Times Co. v. Sullivan, 376 U.S. 254 (1964).

181. *See, e.g.,* Va. Bd. of Pharm. v. Va. Citizens Consumer Council, Inc., 425 U.S. 748 (1976).

182. *See, e.g.,* United States v. Bonin, 932 F.3d 523 (7th Cir. 2019).

183. *See* New York v. Ferber, 458 U.S. 747, 763–64 (1982) (engaging in balancing to determine whether the speech’s “evil to be restricted so overwhelmingly outweighs the expressive interests”); *see also* Young v. Am. Mini Theatres, Inc., 427 U.S. 50, 70 (1976) (plurality opinion) (“[I]t is manifest that society’s interest in protecting this type of expression is of a wholly different, and lesser, magnitude than the interest in untrammelled political debate . . .”).

the First Amendment for two reasons.¹⁸⁴ First, he claims the falsity “can take a form that causes or threatens harm to person or property.”¹⁸⁵ Second, a category of false speech may be unprotected when it lacks the value that justifies First Amendment protection.¹⁸⁶

a. Defamation

Defamation¹⁸⁷ is an example of an exception to protected speech under the First Amendment. In *New York Times Co. v. Sullivan*,¹⁸⁸ the Court held that the First Amendment does not protect knowingly false statements about public officials.¹⁸⁹ In balancing public figures’ interests in maintaining their reputations and protecting high-value First Amendment speech, the Court held that the “actual malice” standard establishes liability for defamation.¹⁹⁰ “Actual malice” determines a person made the statement with knowledge or recklessness as to its falsity.¹⁹¹ The Court justified the high standard by emphasizing that “erroneous statement[s] are] inevitable in free debate.”¹⁹²

The Court noted speech containing factual error or causing reputational injury alone is still within First Amendment protections.¹⁹³ However, in balancing competing considerations of reputational harm and freedom of expression, the Court held that the “actual malice” standard provided adequate “breathing space” to protected speech regarding public persons.¹⁹⁴ Without knowledge or recklessness as to a statement’s falsity, the concurrence emphasized that the press should err on the side of publication to avoid journalistic self-silencing.¹⁹⁵ After *Sullivan*, the decision protecting false speech concerning public officials was extended to public figures.¹⁹⁶

In *Gertz v. Robert Welch, Inc.*,¹⁹⁷ the Court held that ordinary people can recover for defamation when speech is made with at least negligence as to its falsity.¹⁹⁸ In establishing a negligence requirement, the Court determined

184. See Blitz, *supra* note 177, at 77.

185. *Id.*

186. *Id.* at 78.

187. “To prove prima facie defamation, a plaintiff must show . . . (1) a false statement purporting to be fact; (2) publication or communication of that statement to a third person; (3) fault amounting to at least negligence; and (4) damages, or some harm caused to the reputation of the person or entity who is the subject of the statement.” *Defamation*, CORNELL L. SCH. LEGAL INFO. INST., <https://www.law.cornell.edu/wex/defamation> [<https://perma.cc/Q5XL-ZXPU>] (last visited Aug. 31, 2024).

188. 376 U.S. 254 (1964).

189. *Id.* at 279–81.

190. *Id.*

191. *See id.*

192. *Id.* at 271.

193. *Id.* at 272–73.

194. *Id.* at 271–72, 279–81.

195. *See id.* at 299 (Black, J., concurring).

196. *See Curtis Publ’g Co. v. Butts*, 388 U.S. 130, 155 (1967) (extending the “actual malice” standard for defamation to public figures).

197. 418 U.S. 323 (1974).

198. *Id.* at 350.

states can define the appropriate standard for liability.¹⁹⁹ However, the decision noted states may not impose liability for defamation without fault.²⁰⁰ The Court's reasoning similarly acknowledged chilling speech concerns discussed in *Sullivan*.²⁰¹ It emphasized that protecting some falsehoods is imperative to prevent a chilling effect on "speech that matters."²⁰² Ultimately, *Gertz*'s required showing of negligence imposes a similar safeguard to *Sullivan*'s "actual malice" standard to prevent "journalistic or speaker self-silencing."²⁰³

b. *Fraud*

Fraud is another category of false speech that the First Amendment does not shield.²⁰⁴ "The common law of fraud imposes liability on a person who makes a fraudulent misrepresentation" to induce someone to act, "detrimentally and justifiably, in reliance on that material misrepresentation."²⁰⁵ A false statement alone does not subject a [speaker] to fraud liability" unless the speaker intended to and successfully deceived the listener.²⁰⁶ The Court has recognized that the government has a substantial interest in protecting people against fraud.²⁰⁷ As a result, a regulation limiting fraudulent conduct to protect the public can be upheld against a First Amendment challenge.²⁰⁸ In the context of advertisements, the Court has held that the question of fraud may be determined in light of the effect an advertisement would "probably produce on ordinary minds."²⁰⁹

c. *Impersonation of Government Officials*

Another category of false speech that the government is typically permitted to regulate is the impersonation of government officials.²¹⁰ These statutes typically require a specific *act* of impersonation, not only speech, and can require showing a person was misled into action that they would not

199. *Id.* at 347.

200. *Id.*

201. *See id.* at 340–41.

202. *Id.* at 341.

203. Sunstein, *supra* note 22, at 410.

204. *See Illinois ex rel. Madigan v. Telemarketing Assocs., Inc.*, 538 U.S. 600, 612 (2003).

205. VALERIE C. BRANNON, CONG. RSCH. SERV., IF12180, FALSE SPEECH AND THE FIRST AMENDMENT: CONSTITUTIONAL LIMITS ON REGULATING MISINFORMATION (2022), <https://rsreports.congress.gov/product/pdf/IF/IF12180> [<https://perma.cc/D9W4-VFQA>].

206. *Madigan*, 538 U.S. at 620.

207. *United States v. Alvarez*, 567 U.S. 709, 723 (2012) (plurality opinion) ("Where false claims are made to effect a fraud . . . the Government may restrict speech without affronting the First Amendment.").

208. *See, e.g., Donaldson v. Read Mag.*, 333 U.S. 178, 191–92 (1948).

209. *Id.* at 189.

210. *See, e.g., United States v. Lepowitch*, 318 U.S. 702, 704 (1943); *United States v. Bonin*, 932 F.3d 523, 544 (7th Cir. 2019); *see also Chesney & Citron, supra* note 12, at 1791 (noting that impersonation of government official statutes are "largely uncontroversial as a First Amendment matter" (quoting Helen Norton, *Lies to Manipulate, Misappropriate, and Acquire Governmental Power*, in *LAW AND LIES* 143, 170–71 (Austin Sarat ed., 2015))).

have pursued but for the deceitful conduct.²¹¹ Unlike other regulable categories of false speech, the harm from the impersonation of a government official is less tangible.²¹² In justifying restricting this speech, the *Alvarez* plurality noted that there was a compelling government interest in protecting the “integrity of Government processes.”²¹³ Further, the plurality acknowledged that even if the impersonation statute does not require proving a physical, financial, or property harm, it is permissible because of its narrow application.²¹⁴

II. IS A FEDERAL BAN OF POLITICAL ADVERTISEMENT DEEPAKES OF FEDERAL CANDIDATES COMPATIBLE WITH THE FIRST AMENDMENT?

With the growth in accessibility to and sophistication of AI, various regulatory efforts have emerged. Legislation limiting deepfakes in political advertisements is content based and likely subject to strict scrutiny review.²¹⁵ Although this legislation is viewpoint neutral, it involves the restriction of political campaign subject matter.²¹⁶ Moreover, assessing the media’s falsity necessarily requires evaluating its content.²¹⁷

This part will explore the constitutionality of a federal ban of deepfakes depicting federal candidates in political advertisements, as proposed by the Protect Elections from Deceptive AI Act. This bill seeks to ban the distribution of materially deceptive AI-generated audio or visual media relating to candidates for federal office with the purpose of influencing an election or soliciting funds.²¹⁸ Based on case precedent, it is uncertain whether legislation prohibiting knowing AI-generated deepfakes in election advertisements will be constitutional under the First Amendment.²¹⁹ Part II.A examines opponents’ arguments that the bill will fail to be constitutional under the First Amendment. Part II.B considers proponents’ arguments that the proposed ban is constitutional.

A. *The Proposed Ban Is Unconstitutional Under the First Amendment*

This section examines arguments against implementing a ban of federal candidate deepfakes in political advertisements. Part II.A.1 considers

211. *Lepowitch*, 318 U.S. at 704.

212. See Chesney & Citron, *supra* note 12, at 1791.

213. *United States v. Alvarez*, 567 U.S. 709, 721 (2012) (plurality opinion).

214. *Id.* (“[T]he statute is itself confined to ‘maintain[ing] the general good repute and dignity of . . . government . . . service itself.’” (citation omitted)).

215. See *Citizens United v. FEC*, 558 U.S. 310, 340 (2010) (“Laws that burden political speech are ‘subject to strict scrutiny.’” (citation omitted)).

216. *Id.*

217. See VICTORIA L. KILLION, CONG. RSCH. SERV., IF12308, FREE SPEECH: WHEN AND WHY CONTENT-BASED LAWS ARE PRESUMPTIVELY UNCONSTITUTIONAL (2023), <https://crs.reports.congress.gov/product/pdf/IF/IF12308> [<https://perma.cc/B4GW-8X3A>].

218. Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

219. Weiner & Norden, *supra* note 91.

constitutional barriers to regulating deepfakes because they are false speech. Part II.A.2 analyzes issues that the bill raises by limiting political speech. Part II.A.3 explores whether the proposed ban is sufficiently narrowly tailored to pass strict scrutiny. Part II.A.4 evaluates concerns of partisan enforcement to censor speakers.

1. Falsity Alone: The Ban Under *Alvarez*

Those who oppose deepfake regulation emphasize that the nature of the falsity in deepfakes is not akin to constitutionally regulable false speech.²²⁰ In *Alvarez*, the Court established that the government may regulate specific categories of false speech if they cause a “legally cognizable” harm to identifiable victims.²²¹ The risk that deepfakes pose to voters is difficult to ascertain, and it is not harm that the Court recognizes warrants regulation.²²² Adversaries to deepfake regulation also argue that AI-generated deepfakes are no different than typical manipulated media that has historically remained unregulated.²²³ Based on the marketplace of ideas, the response to deepfake political advertisements should be more speech, not government-enforced silence.²²⁴ Challengers to legislation claim that allowing regulation of falsity in this context would impermissibly expand the government’s authority to limit everyday lies.²²⁵ Therefore, opponents claim that deepfake legislation is unconstitutional.²²⁶

The Court’s decision in *Alvarez* determined that the First Amendment protects false speech alone.²²⁷ As a result, scholars and commentators argue that because the Supreme Court held that the First Amendment protects verbal lies, it should also protect visual lies in deepfakes.²²⁸ The plurality in *Alvarez* noted there are categories of false speech that the government may regulate, such as defamation, fraud, and impersonation.²²⁹ In explaining

220. See Alexandra Tashman, Note, “Malicious Deepfakes”—How California’s A.B. 730 Tries (and Fails) to Address the Internet’s Burgeoning Political Crisis, 54 LOY. L.A. L. REV. 1391, 1412 (2021); see also *AI and the Future of Our Elections*, *supra* note 141, at 14 (prepared statement of Ari Cohn, Counsel, TechFreedom).

221. *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

222. See Tashman, *supra* note 220, at 1412 (stating that political deepfakes do not fall into a false speech category that the Court has recognized may be restricted).

223. See Daniel Immerwahr, *What the Doomsayers Get Wrong About Deepfakes*, NEW YORKER (Nov. 13, 2023), <https://www.newyorker.com/magazine/2023/11/20/a-history-of-fake-things-on-the-internet-walter-j-scheirer-book-review> [<https://perma.cc/67Q6-ABZ2>]; see also Alex Baiocco, *Political “Deepfake” Laws Threaten Freedom of Expression*, INST. FOR FREE SPEECH (Jan. 5, 2022), <https://www.ifs.org/research/political-deepfake-laws-threaten-freedom-of-expression/> [<https://perma.cc/AN67-5MEN>].

224. See Baiocco, *supra* note 223.

225. See SEANA VALENTINE SHIFFRIN, *SPEECH MATTERS: ON LYING, MORALITY, AND THE LAW* 150 (2014).

226. See Baiocco, *supra* note 223; see also Tashman, *supra* note 220, at 1408.

227. *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

228. Nina I. Brown, *Deepfakes and the Weaponization of Disinformation*, 23 VA. J.L. & TECH. 1, 35 (2020) (“But the fact that a deepfake is a manufactured video and may contain falsehoods . . . does not weaken its First Amendment protection.”).

229. See *supra* Part I.C.4.

these examples, the Court noted when falsity is unprotected, there is a “legally cognizable harm” associated with the speech.²³⁰ In the context of deepfakes, the harm is more amorphous and not of the sort that typically justifies regulation of false speech.²³¹

Although *Alvarez* addressed circumstances when limitation of speech caused more intangible harms,²³² critics argue they are distinguishable from the harms deepfakes cause.²³³ The government’s argument in *Alvarez* cited to laws punishing impersonation of a government official as speech the Court believed could be constitutionally limited.²³⁴ The Court’s justification for restricting speech impersonating a government official was because of the harm this false speech has on the “general good repute and dignity of . . . government . . . service itself.”²³⁵

Applying the Court’s reasoning for impersonation to deepfakes can stretch what is understood to be a “legally cognizable harm.”²³⁶ As a result, scholars like Professor Blitz argue permitting restriction of deepfakes based upon their “legally cognizable harm,” such as misleading voters or diminished trust in government, can open the door to an impermissible increase in government speech restriction.²³⁷ Professor Blitz notes that if falsehoods are provided “any meaningful First Amendment protection,” the protection cannot be withdrawn whenever harm exists because falsehoods always cause harm.²³⁸ As a result, allowing deepfake regulation could result in the overregulation of false speech.²³⁹

Another category of false speech that the First Amendment does not protect is defamation. As mentioned, for public officials or public figures to succeed in a defamation suit, they must establish that a speaker acted with “actual malice” or “reckless disregard of the truth.”²⁴⁰ Similarly to defamation, political advertisement deepfakes are a knowing falsehood that can cause reputational harm to the individual depicted.²⁴¹ Although critics acknowledge that the proposed ban could correctly prohibit some defamatory deepfakes of federal candidates, they argue it extends too far.²⁴² The bill broadly applies to any AI-generated media of a federal candidate in a political

230. *Alvarez*, 567 U.S. at 719.

231. See Marc Jonathan Blitz, *Deepfakes and Other Non-testimonial Falsehoods: When Is Belief Manipulation (Not) First Amendment Speech?*, 23 YALE J.L. & TECH. 160, 220 (2020).

232. See *supra* note 212 and accompanying text.

233. See Tashman, *supra* note 220, at 1412; see also Blitz, *supra* note 231, at 220.

234. See *supra* note 212 and accompanying text.

235. *Alvarez*, 567 U.S. at 721 (plurality opinion) (quoting *United States v. Lepowitch*, 318 U.S. 702, 704 (1943)).

236. See Blitz, *supra* note 231, at 220.

237. See *id.*

238. See *id.*

239. See *id.*

240. See *supra* notes 190–91 and accompanying text.

241. See Jessica Ice, Note, *Defamatory Political Deepfakes and the First Amendment*, 70 CASE W. RESV. L. REV. 417, 433–34 (2019).

242. See *AI and the Future of Our Elections*, *supra* note 141, at 13–14 (prepared statement of Ari Cohn, Counsel, TechFreedom).

advertisement that intends to influence an election or solicit funds.²⁴³ Moreover, it does not require injury to reputation, which is essential to a defamation claim.²⁴⁴ Consequently, opponents claim that a broad prohibition of political advertisement deepfakes could wrongly stifle protected false speech.²⁴⁵

Fraud is another type of false speech that the First Amendment does not protect. Fraud requires a knowing falsehood that is material, relied upon by the victim, and causes actual injury.²⁴⁶ Similarly to defamation, there are political advertisement deepfakes that could qualify as fraud.²⁴⁷ For example, a voter could wrongfully rely upon a deepfake advertisement depicting a candidate soliciting funds for their campaign and donate to the campaign based on this belief. However, the bill does not require reliance or intent to induce reliance, which is a crucial facet of the common law of fraud.²⁴⁸ Therefore, critics assert that the ban could again wrongfully include speech that may not be constitutionally prohibited.²⁴⁹

Following the *Alvarez* decision, lower courts struck down state legislation that sought to limit lies during campaigns.²⁵⁰ In *Susan B. Anthony List v. Driehaus*,²⁵¹ the U.S. Court of Appeals for the Sixth Circuit struck down as unconstitutional an Ohio law that banned dissemination of knowing or reckless falsehoods about a political candidate in campaign materials.²⁵² In its decision, the court stated that the statute violated the First Amendment partially because it targeted falsity alone.²⁵³ The court cited *Alvarez* and reasoned that “[e]ven false speech receives some constitutional protection.”²⁵⁴ Critics of deepfake regulation argue that it similarly seeks to target generally false—and therefore protected—speech.²⁵⁵ Although unprotected speech is within the ban’s ambit, critics claim that it applies too broadly to falsity alone and is therefore unconstitutional.²⁵⁶

243. Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

244. See *AI and the Future of Our Elections*, *supra* note 141, at 13–14 (prepared statement of Ari Cohn, Counsel, TechFreedom).

245. See *id.* at 18.

246. *Id.* at 4.

247. *Id.* at 13.

248. *Id.* at 13 & n.68 (citing RESTATEMENT (SECOND) OF TORTS § 525 (AM. L. INST. 1977) (“One who fraudulently makes a misrepresentation of fact . . . for the purpose of inducing another to act or to refrain from action in reliance upon it, is subject to liability to the other in deceit for pecuniary loss caused to him by his justifiable reliance upon the misrepresentation.”)).

249. See *id.* at 7–8.

250. See, e.g., *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 476 (6th Cir. 2016); *Commonwealth v. Lucas*, 34 N.E.3d 1242, 1257 (Mass. 2015); *Rickert v. State Pub. Disclosure Comm’n*, 168 P.3d 826, 829–31 (Wash. 2007).

251. 814 F.3d 466 (6th Cir. 2016).

252. *Id.* at 473 (“Ohio’s laws reach not only defamatory and fraudulent remarks, but *all* false speech . . . even that which may not be material, negative, defamatory, or libelous.”).

253. *Id.* at 472.

254. See *id.* at 473.

255. See Brown, *supra* note 228, at 34–35; see also Tashman, *supra* note 220, at 1415.

256. See *AI and the Future of Our Elections*, *supra* note 141, at 13–14 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also Brown, *supra* note 228, at 34–35.

Those who oppose legislation argue that although deepfake technology is new, questions as to whether it is constitutionally regulable have already been answered.²⁵⁷ Historically, the First Amendment generally protects doctored media, which has been used in campaign materials for years.²⁵⁸ Deepfakes, challengers argue, are no different.²⁵⁹ The Supreme Court has held that, notwithstanding the difficulty of applying the Constitution to advancing technology, the basic principles of freedom of speech remain unchanged even when a new medium for communication appears.²⁶⁰ Thus, despite the novelty of AI technology, regulatory opponents claim that it is afforded the same First Amendment protections as garden-variety manipulated media in political advertisements.²⁶¹

Moreover, the marketplace of ideas theory informs critics' assessment of the proposed ban. The Court has held that even a false statement can be a valuable contribution to the marketplace of ideas because it creates a "clearer" and "livelier impression" of truth based upon its "collision with error."²⁶² As a result, the preferred First Amendment remedy to false speech is "more speech, not enforced silence."²⁶³ Therefore, adversaries claim legislation limiting deepfakes can interfere with the marketplace of ideas.²⁶⁴ Despite the confusion deepfakes might cause, under the First Amendment, the proper solution is more debate, not prohibition.²⁶⁵ Positive First Amendment theory indicates that functioning of the marketplace of ideas is dependent on public debate, without government interference.²⁶⁶ Thus, in response to deepfakes, the media, campaigns, and public should counteract the falsity with facts and truth instead of government-enforced silence.²⁶⁷

Prohibiting deepfakes runs the risk of expanding the arena in which the government can regulate false speech.²⁶⁸ The plurality in *Alvarez* expressed

257. See MATTHEW FEENEY, DEEFAKE LAWS RISK CREATING MORE PROBLEMS THAN THEY SOLVE, REGUL. TRANSPARENCY PROJECT 1–2 (2021), <https://regproject.org/wp-content/uploads/Paper-Deepfake-Laws-Risk-Creating-More-Problems-Than-They-Solve.pdf> [<https://perma.cc/8P5P-LGL4>].

258. For example, in 1950, Senator Joseph R. McCarthy's allies circulated doctored photographs of Senator Millard E. Tydings meeting with the U.S. Communist Party leader. See *id.* Senator Tydings lost his election. See *id.* at 2–3. Recently, in 2016, a false photo circulated of Senator Marco Rubio shaking hands with President Barack Obama. See *id.* at 3.

259. See Letter from Christopher Anders, Jenna Leventoff & Brian Hauss, ACLU, to Dara S. Lindenbaum, Chair, FEC 3–4 (Nov. 6, 2023), https://www.nextgov.com/media/aclu_fec_a_i_letter.pdf [<https://perma.cc/EN8Q-DUAH>].

260. See *id.* at 3.

261. See *id.* at 3–4; FEENEY, *supra* note 257, at 2.

262. See *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279 n.19 (1964).

263. See *Whitney v. California*, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring).

264. See *Brown*, *supra* note 228, at 35–36; see also Baiocco, *supra* note 223 ("As the Supreme Court has reiterated time and again, the answer to false and deceptive speech is counterspeech.").

265. See *AI and the Future of Our Elections*, *supra* note 141, at 18 (prepared statement of Ari Cohn, Counsel, TechFreedom).

266. See *United States v. Alvarez*, 567 U.S. 709, 727–28 (2012) (plurality opinion).

267. *AI and the Future of Our Elections*, *supra* note 141, at 10–11 (prepared statement of Ari Cohn, Counsel, TechFreedom).

268. See *Blitz*, *supra* note 231, at 220.

concern that banning lies about military honors would prompt a broader limitation of lies “without a judicial backstop.”²⁶⁹ The Court in *Alvarez* noted that requiring a “legally cognizable harm” to identifiable victims mitigates concerns of government overregulation of false speech.²⁷⁰ This reasoning likewise applies to criticism of the Protect Elections from Deceptive AI Act. Allowing false speech regulation in additional contexts runs the risk of impermissibly expanding government prohibition of lies in everyday life.²⁷¹

Therefore, critics of the proposed ban argue that it fails to establish a “legally cognizable harm” that the Court recognizes allows for constitutional regulation of false speech.²⁷² They claim deepfakes are akin to typical deceptive media in campaign advertisements and do not establish a novel threat to democracy.²⁷³ Thus, in lieu of legislation, opponents of a ban advocate for more speech in response to political advertisement deepfakes.²⁷⁴

2. Unconstitutional Limitation of Political Speech

The proposed bill raises constitutional questions based not only on targeting false speech, but also based on its potential to chill highly protected political speech.²⁷⁵ Central to the First Amendment is the open discussion of candidates so that voters can make informed decisions.²⁷⁶ The Court has emphasized the importance of the First Amendment’s protection of political campaign speech.²⁷⁷ Critics of a ban underscore the value that political speech—even false political speech—provides to the electorate.²⁷⁸ Therefore, opponents note that its impact is likely to encroach upon First Amendment protected political speech.²⁷⁹

During a Senate Committee hearing on “AI and the Future of Our Elections,” First Amendment attorney Ari Cohn explained different

269. See Daniela C. Manzi, Note, *Managing the Misinformation Marketplace: The First Amendment and the Fight Against Fake News*, 87 FORDHAM L. REV. 2623, 2635 (2019) (citing *Alvarez*, 567 U.S. at 723 (plurality opinion)).

270. See *Alvarez*, 567 U.S. at 723.

271. See SHIFFRIN, *supra* note 225, at 150; see also Blitz, *supra* note 231, at 220.

272. See *AI and the Future of Our Elections*, *supra* note 141, at 4–5, 13–14 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also Tashman, *supra* note 220, at 1412.

273. See *supra* notes 257–61 and accompanying text.

274. See *supra* notes 262–67 and accompanying text.

275. See FEENEY, *supra* note 257, at 11.

276. See CHEMERINSKY, *supra* note 95, at 811.

277. See *Monitor Patriot Co. v. Roy*, 401 U.S. 265, 272 (1971).

278. See Baiocco, *supra* note 223 (“[P]unishing false or misleading political speech will inevitably suppress political speech generally and do more harm than good.”); see also Chesney & Citron, *supra* note 12, at 1789–90 (“[A ban] might inhibit engagement in political discourse specifically, and in democratic culture more generally.”).

279. See Baiocco, *supra* note 223; Jacob Fischler, *U.S. Senate Panel Weighs Free Speech and Deep Fakes in AI Campaign Ads*, MINN. REFORMER (Sept. 27, 2023, 10:08 PM), <https://minnesotareformer.com/2023/09/27/u-s-senate-panel-weighs-free-speech-and-deep-fakes-in-ai-campaign-ads/> [<https://perma.cc/E5CB-4739>].

categories of election-related speech and their associated risks.²⁸⁰ He noted that the Court has made a distinction between harm to the electoral process and harm concerning electoral substance.²⁸¹

Regulations that limit speech based on its damage to the electoral process are less likely to run afoul of the First Amendment.²⁸² Cohn notes that although the government is typically permitted to regulate these falsehoods, a federal deepfake political advertisement ban is unlikely to fall into this category.²⁸³

Conversely, legislation targeting harm to electoral substance is more uncertain regarding its constitutionality.²⁸⁴ Legislation implicating electoral conversation speech seeks to regulate based on the speech's content.²⁸⁵ Cohn explains that there is debate among courts as to whether the government has a compelling interest in protecting the public from false speech during campaigns.²⁸⁶ However, even when a compelling government interest is identified, regulations can fail to be sufficiently narrowly tailored to overcome strict scrutiny.²⁸⁷ Thus, when legislation targets political speech based upon its content, it is more likely to be found unconstitutional.²⁸⁸

Cohn argues that the proposed bill broadly regulates electoral substance, and will likely fail strict scrutiny.²⁸⁹ To pass strict scrutiny, a compelling interest must be identified.²⁹⁰ Regulating deepfakes, some claim, does not identify a sufficiently compelling government interest.²⁹¹ For example, some point to the Supreme Court's prior holding that the fear a voter might make an ill-advised choice is not a compelling interest.²⁹² Ultimately, the risk of deepfake political advertisements is that voters will rely upon them when casting their vote. Fear that a voter might rely upon manipulated media in an

280. *AI and the Future of Our Elections*, *supra* note 141, at 6–8 (prepared statement of Ari Cohn, Counsel, TechFreedom).

281. *See id.*

282. *See id.* at 6.

283. *Id.* at 7.

284. *Id.*

285. *Id.* at 6.

286. *Id.* at 7–8.

287. *Id.* at 8; *see, e.g., Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 473–74 (6th Cir. 2016).

288. *See AI and the Future of Our Elections*, *supra* note 141, at 7–8 (prepared statement of Ari Cohn, Counsel, TechFreedom); *see, e.g., Driehaus*, 814 F.3d at 473–74.

289. *See AI and the Future of Our Elections*, *supra* note 141, at 7, 14 (prepared statement of Ari Cohn, Counsel, TechFreedom).

290. CHEMERINSKY, *supra* note 95, at 468.

291. *See Chesney & Citron*, *supra* note 12, at 1803 (“[C]ourts therefore have struck down periodic attempts to ban election-related lies. The entry of deep fakes into the mix may not change that result.” (citation omitted)); *see also Tashman*, *supra* note 220, at 1412 (“Even if this interest was compelling, a court would still have to assess whether the rationale was *compelling enough* to justify the rather heavy burden . . . place[d] on free speech.”).

292. *Brown v. Hartlage*, 456 U.S. 45, 60 (1982).

advertisement, opponents argue, fails to establish a compelling interest and is at odds with self-governance values.²⁹³

Punishing falsity in the political space challenges the ability of the public to engage in “vigorous expression of views . . . the First Amendment seeks to guarantee.”²⁹⁴ Through legislation, the government is obstructing the free flow of information during campaigns that is essential to the democratic process.²⁹⁵ As a result, permitting the government to limit deepfakes could contradict the theory underlying the First Amendment.

Thus, opponents claim that the bill is unconstitutional because it limits highly protected political speech. They assert that the proposed ban fails to establish a compelling government interest to justify restricting speech.²⁹⁶ Regulating deepfakes in this context interferes with the citizenry’s ability to engage in open debate.²⁹⁷ Therefore, opponents argue that deepfake regulation violates the First Amendment.

3. The Proposed Ban Is Not Narrowly Tailored

Ari Cohn argues that even if the government can establish a compelling interest in supporting the bill, it is unlikely to be sufficiently narrowly tailored to pass strict scrutiny.²⁹⁸ Any regulation that targets deepfakes in political advertisements will likely be subject to strict scrutiny and must be narrowly tailored to overcome constitutional challenges.²⁹⁹ To be upheld, the legislation must not be overinclusive or underinclusive, and the government must demonstrate that the speech restriction is “actually necessary” to achieve a compelling interest.³⁰⁰ The proposed ban, Cohn argues, fails to satisfy these requirements.³⁰¹

In his Senate testimony, Cohn asserted that the Protect Elections from Deceptive AI Act is both underinclusive and overinclusive.³⁰² He states that it is underinclusive because it only targets AI-generated deceptive media.³⁰³ Cohn claims that by failing to target types of non-AI-generated misleading media, the ban does not adequately address the harm to voters.³⁰⁴ Critics of the legislation argue that if the compelling interest is the injury to voters from

293. See Tashman, *supra* note 220, at 1411 (“Based on this precedent, regulating altered media, even if that media could manipulate voters into making an ‘ill-advised choice,’ would not satisfy a compelling government purpose.”).

294. *United States v. Alvarez*, 567 U.S. 709, 718 (2012) (plurality opinion).

295. See *AI and the Future of Our Elections*, *supra* note 141, at 3 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also Brown, *supra* note 228, at 34–35.

296. See *supra* notes 291–93 and accompanying text.

297. See *supra* notes 294–95 and accompanying text.

298. See *AI and the Future of Our Elections*, *supra* note 141, at 7–8, 16–18 (prepared statement of Ari Cohn, Counsel, TechFreedom).

299. *Id.*; see also Tashman, *supra* note 220, at 1409.

300. See *supra* notes 101–05 and accompanying text.

301. *AI and the Future of Our Elections*, *supra* note 141, at 16–18 (prepared statement of Ari Cohn, Counsel, TechFreedom).

302. *Id.* at 16–18.

303. *Id.* at 18.

304. *Id.*

consuming deceptive media, it should include other non-AI deceptively edited media.³⁰⁵ They argue that using simpler technology to fabricate media causes the same damage to democracy as AI-generated media.³⁰⁶ Thus, if the legislation seeks to protect voters, it should encompass all deceptive media in political advertisements.³⁰⁷

Critics also assert that the ban is overinclusive and will target and chill valued political speech.³⁰⁸ In Cohn's testimony, he criticized the breadth of the legislation.³⁰⁹ He noted that the ban is overbroad in that it prohibits all individuals from knowingly disseminating deceptive AI-generated media to anyone, including to family and friends.³¹⁰ Moreover, he notes that the bill does not have a limited timeline; its application begins when an individual announces their candidacy.³¹¹ Therefore, he claims the proposed ban is overbroad in its application.

Opponents of deepfake regulatory efforts claim that crafting a statute that is sufficiently narrowly tailored to pass strict scrutiny is difficult to achieve.³¹² Although the proposed legislation could target a compelling interest, the First Amendment requires the government's chosen restriction to be necessary to achieve the interest and prevent the injury.³¹³ Those who oppose the bill have adopted and referenced a recent Texas Court of Appeals' decision.³¹⁴

In May 2023, a Texas Court of Appeals struck down a state statute that criminalized the creation and dissemination of deepfake political advertisements.³¹⁵ Although the court acknowledged that the government established a compelling interest,³¹⁶ the statute was not narrowly tailored.³¹⁷ The decision stated that when free speech values are at stake, the rationale

305. *Id.*; see also *AI and the Future of Our Elections*, *supra* note 141, at 6 (prepared statement of Neil Chilson, Senior Research Fellow at The Center for Growth and Opportunity at Utah State University) ("If the bill's backers truly worry about altered political communications about candidates, why limit this to AI tools?").

306. *AI and the Future of Our Elections*, *supra* note 141, at 18 (prepared statement of Ari Cohn, Counsel, TechFreedom) ("Deceptively edited media produced manually does not merely affect the purported government interest in a comparable way; it affects it in the *same* way.").

307. *See id.*

308. *See id.* at 16–17; see also FEENEY, *supra* note 257, at 11.

309. *See AI and the Future of Our Elections*, *supra* note 141, at 16–17 (prepared statement of Ari Cohn, Counsel, TechFreedom).

310. *Id.*

311. *Id.* at 17.

312. *See Brown*, *supra* note 228, at 45–46; see also Tashman, *supra* note 220, at 1412–16 (concluding that California's deepfake law is not narrowly tailored).

313. CHEMERINSKY, *supra* note 95, at 591.

314. *See AI and the Future of Our Elections*, *supra* note 141, at 8–9 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also *Ex parte Stafford*, 667 S.W.3d 517, 525–26 (Tex. Ct. App. 2023).

315. *Ex parte Stafford*, 667 S.W.3d at 532.

316. *Id.* at 525 ("[P]romoting honest discourse and preventing misinformation in the political arena are compelling state interests.").

317. *Id.* at 525–26.

for the statute must be greater than speculation of harm.³¹⁸ Common sense alone, the court held, does not satisfy the First Amendment burden to justify restricting speech.³¹⁹ Therefore, the law was not the least restrictive means to avoid the harm because existing statutes addressed misrepresentations in election communications.³²⁰ Ultimately, the court wrote, counterspeech would provide a less speech restrictive solution.³²¹

Opponents of deepfake regulation have embraced the reasoning in *Stafford* and argue that is not necessary to regulate deepfakes.³²² Despite their anticipated dangers, adversaries to legislation claim deepfakes have yet to produce a meaningful impact on an election.³²³ Further, it is difficult to support the claim that a deepfake advertisement influenced an election.³²⁴ Instead, they argue, there are less-restrictive alternatives that can achieve the government's goal. One suggestion is using disclosures to indicate that the advertisement was created with AI technology or that it does not depict real events.³²⁵ Another proposed response is using public education in digital literacy to counteract the harm.³²⁶ Lastly, some recommend requiring candidates and campaigns to pledge to refrain from using AI to deceive voters as a possible solution.³²⁷ Ultimately, critics conclude that the government fails to demonstrate that the ban is sufficiently narrowly tailored.

4. Distrust in Government to Enforce the Ban Apolitically

Those who oppose legislation targeting deepfakes premise their arguments on negative theory or distrust in government to regulate neutrally.³²⁸ Permitting regulation of deepfakes, they argue, runs the risk of manipulation

318. *Id.* at 526.

319. *Id.* at 528.

320. *Id.* at 527 (citing existing statutes such as TEX. ELEC. CODE ANN. § 255.005 (West 2023) (political candidate misrepresentations), TEX. PENAL CODE ANN. § 33.07 (West 2011) (online impersonation), and TEX. CIV. PRAC. & REM. CODE ANN. § 73.0001 (West 2023) (false statement state tort claims)).

321. *Id.* at 258.

322. See *AI and the Future of Our Elections*, *supra* note 141, at 8–9 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also Immerwahr, *supra* note 223.

323. *AI and the Future of Our Elections*, *supra* note 141, at 8–9 (prepared statement of Ari Cohn, Counsel, TechFreedom).

324. *Id.* at 15 (“It is also unclear as a general matter whether . . . [AI-generated deepfakes] actually influence[] voters’ beliefs, and ultimate voting decisions.”).

325. See Sunstein, *supra* note 22, at 421 (“[T]he best response to a deepfake is disclosure, not censorship.”); see also Weiner & Norden, *supra* note 91.

326. See Stefanie Koperniak, *Fostering Media Literacy in the Age of Deepfakes*, MIT NEWS (Feb. 17, 2022), <https://news.mit.edu/2022/fostering-media-literacy-age-deepfakes-0217> [<https://perma.cc/C2NK-BYTW>].

327. See Geoffrey A. Fowler, *Candidates, Take This AI Election Pledge. Or 2024 Might Break Us.*, WASH. POST (Oct. 26, 2023, 12:36 PM), <https://www.washingtonpost.com/technology/2023/10/26/ai-election-2024-deepfake-pledge/> [<https://perma.cc/RTG5-EZRB>].

328. During the Senate hearing on AI and Elections, Republican Senator Bill Hagerty stated that he did not trust President Biden and Congress to weigh deepfake concerns against free speech rights. Fischler, *supra* note 279.

by government authorities.³²⁹ The consequences of implementing legislation prohibiting deepfakes could result in censorship of speakers, potentially causing more damage to democracy than deepfakes themselves.³³⁰ This is of particular concern when the harm at issue is less tangible.³³¹ The bill seeks to prohibit deepfakes disseminated for the purpose of “influencing an election” or “soliciting funds.”³³² When it is difficult to ascertain a deepfake creator’s intent, the risk of partisan enforcement is more likely.³³³

In Justice Breyer’s *Alvarez* concurrence, he acknowledged the challenges of regulating false political speech.³³⁴ Justice Breyer warned of the risk of the government applying the statute selectively to speakers they do not approve.³³⁵ In discussing the drafting of legislation limiting political lies to satisfy First Amendment scrutiny, Justice Breyer recognized that political falsehoods could cause harm through changing voting behavior.³³⁶ However, he notes that this harm must be balanced with the risk of censorship.³³⁷

Based on this reasoning, criticism of the proposed ban posits that it fails to balance these harms adequately. The risk of censorship and partisan prosecution is particularly relevant when political speech is at issue.³³⁸ Although the proposed bill will limit some injurious deepfakes, opponents claim it will simultaneously introduce additional democratic dangers, such as censorship.³³⁹ As a result, adversaries rely on negative theory arguments to demonstrate that the bill is unconstitutional.

B. *The Proposed Ban Is Constitutional Under the First Amendment*

This section considers arguments favoring a federal prohibition of deepfakes portraying federal candidates in political advertisements. Part II.B.1 explains how the proposed ban does not seek to target “falsity alone” and could pass constitutional scrutiny under the reasoning in *Alvarez*. Part II.B.2 explores why deepfakes present a unique threat to voters and democracy, which justifies limiting otherwise protected political speech.

329. See Chesney & Citron, *supra* note 12, at 1790; *United States v. Alvarez*, 567 U.S. 709, 723 (2012) (plurality opinion) (noting regulating false speech could give the government a “broad censorial power”).

330. See Chesney & Citron, *supra* note 12, at 1790.

331. See William P. Marshall, *False Campaign Speech and the First Amendment*, 153 U. PA. L. REV. 285, 297 (2004) (questioning the extent to which false campaign speech actually harms voters).

332. Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

333. See Marshall, *supra* note 331, at 299.

334. See *Alvarez*, 567 U.S. at 737–38 (Breyer, J., concurring).

335. See *id.*

336. *Id.* at 738.

337. *Id.*

338. See Chesney & Citron, *supra* note 12, at 1789–90; *supra* Part II.A.2.

339. *AI and the Future of Our Elections*, *supra* note 141, at 18 (prepared statement of Ari Cohn, Counsel, TechFreedom) (claiming the proposed ban “provides a weapon with which to silence critics or opponents”).

Part II.B.3 demonstrates that the legislation is sufficiently narrowly tailored to withstand strict scrutiny. Part II.B.4 determines that the statute's narrow application mitigates concerns regarding partisan enforcement.

1. More Than “Falsity Alone”: Fitting the Ban into the *Alvarez* Framework

Proponents of deepfake legislation argue that although it limits false speech, it can be found constitutional under the First Amendment.³⁴⁰ They emphasize that political deepfake regulations do not seek to target deepfakes based only upon their falsity.³⁴¹ Instead, supporters argue that a deepfake's harms are more similar to the categories of false speech that are regulable under the First Amendment.³⁴² Although deceptive media is typically not prohibited in political advertisements, AI-generated deepfakes present novel harms due to their realism and lack of verifiable source material.³⁴³ Further, political deepfakes disrupt the marketplace of ideas because they inhibit voters' ability to assess the source and validity of speech.³⁴⁴ Advocates of regulatory efforts claim that narrowly drafted deepfake legislation can mitigate concerns regarding expanding the government's authority to limit false speech.³⁴⁵ Consequently, they argue that the bill is compatible with the First Amendment.

Based upon the reasoning in *Alvarez*, an outright ban of deepfakes is unlikely to pass constitutional scrutiny.³⁴⁶ However, when deepfakes of federal candidates cause a “legally cognizable harm” to identifiable victims, they could be constitutionally regulable.³⁴⁷ Deepfake political advertisements manipulate voters into believing that a candidate did or said something that never happened, and voting based on that false belief. Thus, advocates claim political deepfake legislation targets harm analogous to other types of false speech that the Court in *Alvarez* held to be constitutionally regulable.³⁴⁸

Supporters of legislation assert that deepfakes cause similar amorphous harm to voters and democracy that impersonation of a public official causes to the operations of government.³⁴⁹ In the *Alvarez* decision, the Court recognized that false speech like impersonation causes harm to the trust and

340. See Press Release, *supra* note 74; see also Green, *supra* note 21, at 1483–84.

341. See Green, *supra* note 21, at 1483–84; see also Blitz, *supra* note 231, at 169–71.

342. See Blitz, *supra* note 231, at 170, 219–20.

343. See Green, *supra* note 21, at 1452.

344. See Chesney & Citron, *supra* note 12, at 1791; see also Helen Norton, (*At Least*) *Thirteen Ways of Looking at Election Lies*, 71 OKLA. L. REV. 117, 131–33 (2018).

345. See Green, *supra* note 21, at 1483; see also Pesetski, *supra* note 52, at 529.

346. Chesney & Citron, *supra* note 12, at 1790.

347. See Green, *supra* note 21, at 1483, 1486 (“It provides the narrow exception the *Alvarez* court required: proof of specific harm to identifiable victims or specific harm by interference with the functioning of a government department.”).

348. *Id.*; see also Pesetski, *supra* note 52, at 514–16 (noting the similarities between deepfake harms and defamation and fraud).

349. See Chesney & Citron, *supra* note 12, at 1791–92; see also Green, *supra* note 21, at 1485.

functioning of government, which justifies its prohibition.³⁵⁰ Political deepfakes similarly interfere with the functioning of democracy because they inhibit voters' engagement as members of the electorate.³⁵¹ The public's ability to make informed decisions among candidates for office is essential to democracy.³⁵² Deepfakes impede this protected right because they deceive voters into relying upon unverifiable falsified media when casting their ballot.³⁵³ Thus, supporters argue that the harm to voters and democracy, though more diffuse, is analogous to constitutionally regulable falsehoods.³⁵⁴

The Court's reasoning in establishing liability for defamation is relevant to determining the constitutionality of regulating deepfakes. In *Sullivan*, the Court held that defamatory speech can be limited when the speaker states the lies with "actual malice" and causes reputational or financial harm to a public official.³⁵⁵ Similarly to defamation, the ban seeks to target a knowing falsehood.³⁵⁶ However, unlike defamation, a deepfake creator's knowledge of falsity is more easily established.³⁵⁷ To create an AI-generated deepfake, an individual must take a series of affirmative steps using a computer to generate a fake image, video, or audio.³⁵⁸

Applying this reasoning to the bill, advocates will likely argue that it is constitutional based on the logic of *Sullivan*.³⁵⁹ When a knowing falsehood causes significant harm to a compelling government interest, it can likely be regulated under the First Amendment.³⁶⁰ Deepfake political advertisements of federal candidates involve a knowing falsehood that causes harm to a compelling interest in protecting voters and democracy.³⁶¹ In this way, proponents claim deepfakes are more akin to the unprotected falsehoods the Court recognizes than to the protected "falsity alone" in *Alvarez*.³⁶²

Those in favor of deepfake legislation also assert that deepfakes are a constitutionally regulable form of fraud.³⁶³ Commentators note that there is

350. See *Alvarez*, 567 U.S. at 721.

351. See Chesney & Citron, *supra* note 12, at 1777–79.

352. See *Buckley v. Valeo*, 424 U.S. 1, 14–15 (1976) ("In a republic where the people are sovereign, the ability of the citizenry to make informed choices among candidates for office is essential . . .").

353. See Lindsey Wilkerson, Note, *Still Waters Run Deep(fakes): The Rising Concerns of "Deepfake" Technology and Its Influence on Democracy and the First Amendment*, 86 MO. L. REV. 407, 410 (2021).

354. See *supra* notes 347–48 and accompanying text.

355. See *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279–80 (1964).

356. See Ice, *supra* note 241, at 433–34.

357. *Id.*

358. *Id.*

359. See Redish & Pereyra, *supra* note 135, at 467–68.

360. *Id.*

361. See Sunstein, *supra* note 22, at 419–21.

362. *Id.*; see also Green, *supra* note 21, at 1481; *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

363. See Green, *supra* note 21, at 1448; see also Robert Weissman & Craig Holman, Petition for Rulemaking to Clarify That the Law Against "Fraudulent Misrepresentation" (52 U.S.C. § 30124) Applies to "Deepfakes" in Campaign Communications 3 (FEC Oct. 11, 2023), <https://sers.fec.gov/fosers/showpdf.htm?docid=423835> [https://perma.cc/TF6G-

a difference between regulating a general falsehood and fraud.³⁶⁴ Falsity is something that is not true. In contrast, fraud is intentional deceit of another to obtain something of value.³⁶⁵ Deepfake political advertisements make campaign material realistically appear as though it comes from a candidate when it does not.³⁶⁶ As a result, deepfakes operate as a form of fraud on the voter because they are communicated deceitfully to gain an advantage based on people buying into the lie.³⁶⁷ Moreover, an AI-generated deepfake is more disruptive to the democratic process than typical campaign falsity for the following reasons.

Although doctored media and lying is not new to political races, supporters of regulating AI-generated deepfakes assert that they present novel problems. It is commonplace for campaigns to use manipulated media in advertisements. For example, in 2020, the Biden campaign ran an advertisement against former President Trump in which portions of Trump's speech were edited to make it appear that Trump called COVID-19 a "hoax."³⁶⁸ The video edited out small portions of Trump's speech and combined them to make it appear as if he said something he did not.³⁶⁹ Similarly, in 2020, Trump's campaign ran an advertisement depicting a doctored image of President Biden.³⁷⁰ Trump's campaign removed a microphone and trees from a photo so it appeared that Biden was hiding in his basement.³⁷¹

However, proponents of regulation note that AI-generated deepfakes are uniquely harmful compared to typical manipulative editing techniques.³⁷² First, there is no genuine image, video, or audio to serve as a fact-check for deepfakes.³⁷³ In the prior examples, voters could view the context of Trump's speech or the Biden photos to determine whether these events happened. For deepfakes, however, there is no way to authenticate the media.³⁷⁴ Second, doctored media has not historically achieved the realistic

4R8U] ("The deceptive deepfake is fraudulent because the deepfaked candidate . . . did not say or do what is depicted . . . and because the deepfake aims to deceive the public.").

364. See Green, *supra* note 21, at 1448; see also Weissman & Holman, *supra* note 363, at 3.

365. See *supra* Part I.C.4.b.

366. See Green, *supra* note 21, at 1448.

367. See Wes Henricksen, *Disinformation and the First Amendment: Fraud on the Public*, 96 ST. JOHN'S L. REV. 543, 556 (2022).

368. *AI and the Future of Our Elections*, *supra* note 141, at 3 (prepared statement of Ari Cohn, Counsel, TechFreedom).

369. Meg Kelly, *Biden Ad Manipulates Video to Slam Trump*, WASH. POST (Mar. 13, 2020, 3:00 AM), <https://www.washingtonpost.com/politics/2020/03/13/biden-ad-manipulates-video-slam-trump/> [<https://perma.cc/XWT3-Z6C9>].

370. Daniel Dale & Andrew Kaczynski, *Fact Check: Trump Ad Edits Out Microphone and Trees From Biden Photo to Make Him Seem Alone in Basement*, CNN (Aug. 5, 2020, 6:12 PM), <https://www.cnn.com/2020/08/05/politics/fact-check-trump-ad-biden-basement-delaw-are-photos-iowa/index.html> [<https://perma.cc/X5KF-TPBP>].

371. *Id.*

372. Sunstein, *supra* note 22, at 422; see also Chesney & Citron, *supra* note 12, at 1757.

373. See Green, *supra* note 21, at 1452.

374. Sunstein, *supra* note 22, at 422.

quality that deepfakes have attained.³⁷⁵ Thus, AI deepfakes are “self-authenticating,” because their realism can provide more credibility than other verbal or visual representations.³⁷⁶

When reviewing legislation that limits false speech, the Court typically favors counterspeech rather than prohibiting speech.³⁷⁷ Based on the marketplace of ideas, false speech can be beneficial in encouraging public debate and the search for truth.³⁷⁸ However, some scholars reason that the marketplace is inoperable when the public is unable to assess the validity of a speaker’s claims.³⁷⁹ As a result, legislation targeting deepfakes is consistent with positive First Amendment theory of encouraging meaningful political debate.³⁸⁰ Those in favor of regulating deceptive political speech argue that legislation ensures that only “real speech” enters the marketplace, thereby ensuring its functioning.³⁸¹

Advocates of legislation argue that limiting deepfakes would not result in overregulation of lies.³⁸² They claim that a narrowly drawn prohibition to the specified harms from deepfakes is compatible with the *Alvarez* opinion.³⁸³ Thus, concerns regarding government overregulation of falsity are unwarranted given the proposed ban’s limited application and narrow tailoring.³⁸⁴ As a result, supporters of the bill likely contend that it will not result in government censorship of speakers.

Supporters of political candidate deepfake regulation claim that although deepfakes are falsehoods, the proposed legislative efforts do not seek to limit deepfakes only because they are false.³⁸⁵ Instead, they argue, the legislation targets a unique “legally cognizable harm” that political deepfakes cause to voters and democracy.³⁸⁶ Specifically, the bill targets the fraudulent nature of AI deepfakes because they deceive voters into acting in reliance on the falsified media.³⁸⁷ Therefore, proponents assert it is constitutional to prohibit using AI-generated deepfakes in political advertisements.

375. See Chesney & Citron, *supra* note 12, at 1756.

376. Sunstein, *supra* note 22, at 422.

377. See, e.g., *Whitney v. California*, 274 U.S. 357, 377 (1927); *United States v. Alvarez*, 567 U.S. 709, 726–27 (2012) (plurality opinion); *Citizens United v. FEC*, 558 U.S. 310, 361 (2010).

378. See *supra* Part I.C.2.b; see also *Alvarez*, 567 U.S. at 733 (Breyer, J., concurring).

379. See Green, *supra* note 21, at 1459–60; see also Chesney & Citron, *supra* note 12, at 1778.

380. Green, *supra* note 21, at 1457–59.

381. See *id.* at 1485.

382. See Pesetski, *supra* note 52, at 529 (“However, the narrow tailoring of any ban of deepfakes would effectively limit the control that the government has over false speech.”).

383. See *id.*

384. See *id.*; *infra* Part II.B.3.

385. See *supra* note 341 and accompanying text.

386. See *United States v. Alvarez*, 567 U.S. 709, 719 (2012) (plurality opinion).

387. See Green, *supra* note 21, at 1448.

2. Justifying Limiting Political Speech

Advocates for deepfake legislation claim that it is constitutional despite limiting political speech. The Court acknowledges that speech is not automatically protected because it is political.³⁸⁸ When the government has identified a compelling interest and a statute is narrowly tailored, laws targeting political speech can be upheld.³⁸⁹ Supporters of political deepfake regulation argue that protecting the electoral process from distortions created by deepfakes is a compelling government interest.³⁹⁰ Absent regulation, deepfakes interfere with self-governance values by inhibiting voters' engagement as members of the electorate.³⁹¹ Therefore, proponents likely argue that though the bill targets political speech, it is constitutional.

The Court recognizes that the government has a compelling interest in upholding the integrity of elections³⁹² and maintaining a stable political system.³⁹³ When legislation targets political speech based on the electoral process, such as protecting voters from fraud or undue influence, it can be upheld.³⁹⁴ However, the proposed deepfake ban seeks to limit substantive electoral speech, or the content of political speech, and is therefore on less certain constitutional footing.³⁹⁵

Following *Alvarez*, several false campaign speech regulations were struck down.³⁹⁶ However, some of the lower courts' decisions acknowledged that the statutes established a compelling government interest in limiting false political speech.³⁹⁷ More specifically, in these cases, the identified interest was in protecting voters from misinformation and preserving the integrity of

388. See *Garrison v. Louisiana*, 379 U.S. 64, 75 (1964) (“That speech is used as a tool for political ends does not automatically bring it under the protective mantle of the Constitution. For the use of the known lie as a tool is at once at odds with the premises of democratic government and with the orderly manner in which economic, social, or political change is to be effected.”).

389. See *supra* notes 140, 155 and accompanying text.

390. See Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a “Post-truth” World*, 64 ST. LOUIS U. L.J. 535, 545–46 (2020); see also Pesetski, *supra* note 52, at 531.

391. Chesney & Citron, *supra* note 12, at 1791.

392. See *Burson v. Freeman*, 504 U.S. 191, 199 (1992).

393. See *Eu v. S.F. Cnty. Democratic Cent. Comm.*, 489 U.S. 214, 226 (1989).

394. See *supra* notes 144–49 and accompanying text.

395. See *AI and the Future of Our Elections*, *supra* note 141, at 14 (prepared statement of Ari Cohn, Counsel, TechFreedom); see also *supra* notes 150–54 and accompanying text.

396. See, e.g., *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 475–76 (6th Cir. 2016); *281 Care Comm. v. Arneson*, 766 F.3d 774, 789 (8th Cir. 2014); *Commonwealth v. Lucas*, 34 N.E.3d 1242, 1257 (Mass. 2015).

397. See *Susan B. Anthony List*, 814 F.3d at 473–74; *281 Care Comm.*, 766 F.3d at 786–87; *Ex parte Stafford*, 667 S.W.3d 517, 525 (Tex. Ct. App. 2023).

elections.³⁹⁸ However, the statutes were unconstitutional because they were not narrowly tailored.³⁹⁹

Applying this reasoning, proponents argue that the risk of deepfakes constitutes a compelling interest that justifies limiting political speech.⁴⁰⁰ Although the Supreme Court has held that a state's fear that a voter might make an ill-advised choice is not a compelling interest, advocates for regulation claim that the threat of deepfakes is more significant than an ill-advised choice.⁴⁰¹ Deepfake political advertisements knowingly spread falsity about candidates that cannot be easily—if ever—authenticated.⁴⁰² The danger of deepfakes is that they cause citizens to lose faith in leaders, policies, and the government itself.⁴⁰³ Regulatory proponents contend that deepfakes will cause voters to question everything that a candidate says that they do not witness in person.⁴⁰⁴ Thus, they claim that prohibiting deepfakes in this context is a compelling government interest because the threat is “existential” to the democratic process itself.⁴⁰⁵

Additionally, some claim that regulating deepfakes is more consistent with the positive First Amendment theory of self-governance than not regulating at all.⁴⁰⁶ Dean Robert M. Chesney and Professor Danielle K. Citron note that deepfakes do not advance democratic values.⁴⁰⁷ Instead, such lies deny the voter the ability to evaluate the credibility of the political speech.⁴⁰⁸ When a voter is unable to authenticate the media depicting a candidate, they cannot engage in the political process that lies at the heart of the First Amendment.⁴⁰⁹ Moreover, allowing deepfakes to deceive voters “removes

398. See *Susan B. Anthony List*, 814 F.3d at 473–74 (recognizing a compelling interest in preserving the integrity of elections and protecting voters from undue influence); *281 Care Comm.*, 766 F.3d at 786–87 (acknowledging a compelling interest in preserving fair elections and preventing fraud on the electorate); *Ex parte Stafford*, 667 S.W.3d at 525 (stating that promoting honest discourse and preventing misinformation in elections are compelling interests).

399. See *Susan B. Anthony List*, 814 F.3d at 474 (“But Ohio’s laws do not meet the second requirement: being narrowly tailored to protect the integrity of Ohio’s elections.”); *281 Care Comm.*, 766 F.3d at 789 (“For this reason, among others, the restriction is neither narrowly tailored nor necessary.”); *Ex parte Stafford*, 667 S.W.3d at 521 (“We conclude . . . the restrictions on protected speech are not narrowly tailored to achieve the compelling interests advanced by the [s]tate.”).

400. See *supra* note 390 and accompanying text.

401. See *supra* note 390 and accompanying text.

402. See *Green*, *supra* note 21, at 1452.

403. See *id.* at 1460–61.

404. See *id.* at 1461.

405. See *id.*; see also Redish & Pereyra, *supra* note 135, at 454.

406. See Pesetski, *supra* note 52, at 511–12 (noting that First Amendment justifications of protecting speech based on self-government fail for deepfakes); see also *Green*, *supra* note 21, at 1457–60 (emphasizing how deepfakes undermine self-government and voter autonomy).

407. See Chesney & Citron, *supra* note 12, at 1777–78.

408. See *id.*

409. *Green*, *supra* note 21, at 1457–58.

the legitimacy and representativeness of democracy.”⁴¹⁰ Thus, those favoring deepfake legislation emphasize that failing to limit deepfakes will harm voters and erode self-government.⁴¹¹

The senators that introduced the bill argue that limiting deepfake political advertisements is imperative due to the impact on democracy.⁴¹² Despite restricting political speech, supporters assert that the ban identifies a compelling interest in protecting voters and elections by ensuring the public has accurate information to inform their vote.⁴¹³ Based on this reasoning, despite implicating political speech, those in favor of the bill claim it does not violate the First Amendment.

3. The Proposed Ban Is Narrowly Tailored

The legislators that introduced the bill assert that it is constitutional.⁴¹⁴ First, the proposed ban seeks to target a narrow subset of campaign speech: AI-generated deepfakes of federal candidates in political advertisements.⁴¹⁵ Second, the legislation contains adequate safeguards so that it is not unduly underinclusive or overinclusive.⁴¹⁶ Lastly, it is necessary to curb the threat of deepfakes.⁴¹⁷

Advocates of the ban will likely claim that it is sufficiently narrowly tailored. For the legislation to withstand constitutional scrutiny, it must target a limited subset of campaign speech.⁴¹⁸ The ban seeks to prohibit deepfakes of political candidates that are materially deceptive, AI-generated, and made with the purpose of influencing elections or soliciting funds.⁴¹⁹ Therefore, its narrow focus ensures it does not broadly, or unconstitutionally, limit false speech in political advertisements.⁴²⁰

Moreover, the proposed legislation’s mens rea requirements appear to avoid encroaching upon speech protected by the First Amendment.⁴²¹ The bill includes a knowledge requirement as to the distribution of materially deceptive AI-generated media, similar to the actual malice standard in

410. See Staci Lieffring, Note, *First Amendment and the Right to Lie: Regulating Knowingly False Campaign Speech After United States v. Alvarez*, 97 MINN. L. REV. 1047, 1064 (2013).

411. See Pesetski, *supra* note 52, at 511–12; see also Green, *supra* note 21, at 1457–60.

412. See Press Release, *supra* note 74.

413. See Green, *supra* note 21, at 1461; *supra* note 390 and accompanying text.

414. See Press Release, *supra* note 74.

415. See *id.*

416. See Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. (2024); see also Green, *supra* note 21, at 1486 (noting that “a narrow law targeting counterfeited candidate speech produced and distributed with the knowing intent to confuse voters and disrupt elections should survive First Amendment scrutiny”); O’Donnell, *supra* note 53, at 724 (“[A] law prohibiting the knowing distribution of malicious deepfakes . . . could conceivably survive a First Amendment challenge.”).

417. See Green, *supra* note 21, at 1486 (emphasizing that “the harm [that the legislation] seeks to prevent is democratically existential”).

418. See Green, *supra* note 21, at 1486; see also Pesetski, *supra* note 52, at 514, 517.

419. See S. 2770, § 2.

420. See Green, *supra* note 21, at 1486; see also Pesetski *supra* note 52, at 514, 517.

421. See Green, *supra* note 21, at 1486.

defamation.⁴²² In addition, it requires an actor to distribute a deepfake with the purpose of influencing an election or soliciting funds.⁴²³ Supporters will likely assert that the layered mens rea accountability provides sufficient safeguards to avoid being unduly overinclusive or underinclusive.⁴²⁴ Therefore, advocates of the proposed ban will likely contend it adequately balances the harm deepfakes cause to voters against the value of freedom of expression.⁴²⁵

Lastly, advocates of the proposed ban likely claim that the legislation is necessary and less-restrictive alternatives are inadequate.⁴²⁶ Proponents of the Protect Elections from Deceptive AI Act likely assert that the bill ensures that the most harmful political deepfake advertisements are prohibited and cannot continue to circulate and influence voters.⁴²⁷ Those that support deepfake regulation emphasize that less speech restrictive alternatives, such as disclosure, can fail to address the damage from deepfakes.⁴²⁸ For example, some argue that a disclosure regime for deepfake political advertisements runs the risk of further manipulation, including removing disclosures.⁴²⁹ Similarly, media literacy programs, campaign pledges, and counterspeech also allow voters to continue to view deepfake advertisements. Therefore, supporters of the proposed ban likely underscore that these alternatives fail to mitigate the harm to voters because they continue to allow actors to disseminate the deceptive advertisements.⁴³⁰

Ultimately, advocates of the proposed ban claim that it is narrowly tailored because it targets a limited subset of campaign speech.⁴³¹ Further, supporters claim it is necessary because less-restrictive methods fail to address the consequences to voters and democracy.⁴³² Thus, those in favor of the Protect Elections from Deceptive AI Act assert it can withstand constitutional scrutiny.

422. See S. 2770, § 2; cf. *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 279–80 (1964).

423. See S. 2770, § 2.

424. See *supra* note 416 and accompanying text; see also Pesetski, *supra* note 52, at 518 (stating that California’s deepfake election law appears to pass First Amendment scrutiny because it requires an actor to have both knowledge as to the deepfake’s falsity and to act with the intent to confuse voters).

425. See Green, *supra* note 21, at 1469; see also Pesetski, *supra* note 52, at 518 (claiming that California’s deepfake election seems to pass First Amendment scrutiny by avoiding restricting valued political speech and targeting the spread of harmful deepfakes).

426. See *AI and the Future of Our Elections*, *supra* note 141, at 8 (prepared statement of Trevor Potter, President, Campaign Legal Center); see also Weiner & Norden, *supra* note 91 (noting advantages that outright deepfake prohibitions have over disclosure).

427. See *AI and the Future of Our Elections*, *supra* note 141, at 8 (prepared statement of Trevor Potter, President, Campaign Legal Center) (“For the most pernicious forms of AI-based fraud and manipulation, a prohibition would seem most appropriate.”).

428. See Weiner & Norden, *supra* note 91.

429. See *id.*

430. See *supra* notes 426–29 and accompanying text.

431. See *supra* notes 418–20 and accompanying text.

432. See *supra* notes 426–29 and accompanying text.

4. Safeguards Mitigate Partisan Enforcement Concerns

Proponents of the Protect Elections from Deceptive AI Act will likely claim that it provides sufficient safeguards to protect the public from negative theory concerns, such as government censorship. Professor Rebecca Green posits that a statute limiting deceptive political speech can decrease the potential for partisan abuse.⁴³³ Professor Green has proposed legislation limiting “counterfeit campaign speech.”⁴³⁴ She argues that when a statute is not evaluating falsity, but rather the intent of the speaker, partisan motives are less troublesome.⁴³⁵ Professor Green’s evaluation of the constitutionality of this broader prohibition informs whether the Protect Elections from Deceptive AI Act could similarly withstand constitutional scrutiny.

Therefore, applying Professor Green’s logic to the bill, advocates will likely argue that the risk of partisanship and censorship is reduced. The legislation does not require assessing the truth of content, but rather, assessing whether the deceptive material is knowingly disseminated with the intent to influence an election or solicit funds.⁴³⁶ Moreover, those in favor of regulating deepfakes assert that fears of partisan abuse do not justify failing to regulate harms that threaten to dismantle democracy.⁴³⁷ Thus, supporters of the proposed ban claim it is constitutional under the First Amendment and provides adequate defenses against partisan abuse.

III. IS IT CONSTITUTIONAL?: A NARROW SOLUTION TO A LEGALLY COGNIZABLE HARM

Part III.A concludes that the Protect Elections from Deceptive AI Act is constitutional under the First Amendment. Part III.B determines that less speech restrictive solutions fail to address the harms of federal candidate deepfakes in political advertisements. Part III.C offers revisions to the bill to narrow its application and lessen concerns of partisan enforcement.

A. *The Proposed Ban Is Constitutional*

The fractured opinion in *Alvarez* lacks clarity regarding whether legislation limiting false political speech can be constitutionally regulated.⁴³⁸ Evidently, sweeping prohibitions of lies or general falsity in politics are unconstitutional.⁴³⁹ However, current First Amendment doctrine leaves

433. See Green, *supra* note 21, at 1483–86.

434. Professor Green discusses a federal ban of “counterfeit campaign speech,” which she describes as “instances in which political candidates’ identities, actions, words, and images are intentionally faked with the intent to confuse voters and distort democracy.” *Id.* at 1450. Professor Green’s proposal includes deepfakes and extends to other fabricated campaign speech. See *id.* at 1447.

435. See Green, *supra* note 21, at 1483; see also Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

436. See S. 2770.

437. See Green, *supra* note 21, at 1486.

438. See *supra* notes 174–77 and accompanying text; see also *supra* notes 346–48 and accompanying text.

439. See *supra* notes 250–56 and accompanying text.

room for narrowly crafted legislation that targets false political speech that causes “legally cognizable harm” to identifiable victims.⁴⁴⁰

Ultimately, the Protect Elections from Deceptive AI Act is constitutional. It successfully targets a narrow subset of speech by limiting its application to deceptive AI-deepfake imagery, video, and audio of federal candidates in political advertisements.⁴⁴¹ The bill does not impose a broad ban of false political advertisements. Instead, it targets deepfakes’ particularized harm to voters and democracy.⁴⁴² As a result, the ban fits into the exception outlined in *Alvarez* by prohibiting AI-generated deepfakes that falsely and knowingly depict candidates with the intent of influencing elections or soliciting funds.⁴⁴³

Deepfakes in political advertisements present a unique issue when both regulating and not regulating the risk of harm to democracy.⁴⁴⁴ To balance both risks, legislation must promote First Amendment values without prohibiting protected speech.⁴⁴⁵ The fact that political advertisement deepfakes implicate otherwise highly protected political discourse is not a justification for failing to regulate. Instead, the harm that political deepfakes cause to our political process justifies prohibiting them despite the competing First Amendment risk.⁴⁴⁶

When a compelling government interest exists to protect voters from undue influence or fraud or exists to maintain the stability of the political system, the Supreme Court has permitted regulation of political speech.⁴⁴⁷ The need to deter deepfakes is more than a compelling interest; it is intertwined with a primary principle of the First Amendment—ensuring that the electorate is informed.⁴⁴⁸ If voters are unable to discern falsity from reality, they cannot engage meaningfully in public debate.⁴⁴⁹

The proposed ban is narrowly tailored and necessary to curb the harm of deepfakes. It is not unduly overinclusive or underinclusive, and it provides adequate safeguards to avoid infringing on protected speech.⁴⁵⁰ Opponents’ arguments that the legislation is underinclusive because it does not apply to all doctored media is unconvincing.⁴⁵¹ If the bill covered all doctored media, it would likely be struck down for overbreadth because it seeks to limit “falsity alone.”⁴⁵² Although other doctored media can also mislead voters,

440. *See supra* notes 346–48 and accompanying text; *see also supra* note 416 and accompanying text.

441. *See supra* notes 418–20 and accompanying text.

442. *See supra* notes 360–62 and accompanying text.

443. *See supra* notes 347–48 and accompanying text.

444. *See* Redish & Pereyra, *supra* note 135, at 468 (discussing the harms of both regulating and failing to regulate fraudulent political speech).

445. *Id.*

446. *See supra* notes 183, 404–05 and accompanying text.

447. *See supra* notes 392–93 and accompanying text.

448. *See supra* note 352 and accompanying text.

449. *See supra* notes 406–11 and accompanying text.

450. *See supra* Part II.B.3.

451. *See supra* notes 303–07 and accompanying text.

452. *See supra* Part I.C.3.b; *see also supra* note 160 and accompanying text.

AI-generated deepfakes present a new and more disruptive threat to democracy.⁴⁵³

Similarly, opponents' argument that the proposed ban is unduly overinclusive is unavailing.⁴⁵⁴ The legislation's knowledge and intent requirements ensure that it does not impermissibly expand the government's ability to regulate false speech.⁴⁵⁵ To impose liability, the actor must distribute the material with the purpose of "influencing an election" or "soliciting funds."⁴⁵⁶ Absent this intent, the legislation's penalties do not apply.⁴⁵⁷ As in *Sullivan*, requiring knowledge of the falsity ensures that there is adequate "breathing space" to avoid limiting protected speech.⁴⁵⁸

However, the intent requirement is flawed because it lacks specificity required to pass constitutional scrutiny. It is exceedingly difficult to prove that a creator intended for their deepfake to influence an election's outcome. Also, the ban does not incorporate a timeline when the statute is applicable; it begins once someone announces their candidacy.⁴⁵⁹ Part III.C considers potential revisions to establish more definite intent and time requirements for the proposed statute.⁴⁶⁰

Although the recent use of deepfake advertisements has seemingly not yet influenced an election,⁴⁶¹ regulation is actually necessary.⁴⁶² The public should use fact-checking, adding context, and counterspeech to challenge fabricated media when these methods can successfully offset the lie.⁴⁶³ However, for AI-generated deepfakes, less-restrictive alternatives are insufficient to mitigate the danger of deepfakes.⁴⁶⁴ Part III.B explains why other proposed solutions are inadequate.

Failing to pass federal legislation could cause massive disruptions in upcoming elections. For example, campaigns or PACs could create a deepfake shortly before an election, showing an opponent dropping out of the race and endorsing another candidate.⁴⁶⁵ Alternatively, they could produce a deepfake portraying an opposing candidate taking a hardline stance on a controversial issue, such a gun control or abortion. Moreover, without restrictions, savvy campaigns can use PMTs to target voters who are more

453. See *supra* notes 372–76 and accompanying text.

454. See *supra* notes 308–11 and accompanying text.

455. See *supra* notes 422–25 and accompanying text.

456. Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

457. See *id.*

458. See *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 271–72 (1964).

459. See *supra* note 311 and accompanying text.

460. See *infra* Part III.C.

461. See *supra* note 323 and accompanying text.

462. See *supra* notes 426–30 and accompanying text.

463. See *supra* notes 265–67 and accompanying text.

464. See *supra* notes 426–30 and accompanying text.

465. See Peter Suci, *There Is Now a Deep Fake Video of Ron DeSantis Dropping Out of the 2024 Race*, FORBES (Sept. 2, 2023, 10:50 PM), <https://www.forbes.com/sites/petersuci/2023/09/02/there-is-now-a-deep-fake-video-of-ron-desantis-dropping-out-of-the-2024-race/> [https://perma.cc/95VU-6FDW] (discussing a deepfake of Ron DeSantis dropping out of the 2024 presidential race).

susceptible to believing deepfakes are real with advertisements.⁴⁶⁶ Voters can rely on information in advertisements, which in turn can influence their voting decisions.⁴⁶⁷ Therefore, the fallout from these hypothetical deepfake advertisements has the power to change the result of an election and should not be overlooked.

Lastly, opponents' concerns regarding the government's ability to regulate neutrally are valid.⁴⁶⁸ Partisan enforcement is an important consideration, particularly when it comes to regulating false political speech.⁴⁶⁹ Although the statute mitigates these harms by assessing the deepfake creator's intent—rather than the truth of the content⁴⁷⁰—there are other measures that legislators should consider. Part III.C suggests that enforcement by a neutral government agency could quell fears of government censorship.⁴⁷¹

B. Less Speech Restrictive Alternatives Are Ineffective

Those opposing the proposed ban assert that it is not “actually necessary.”⁴⁷² They assert that when the government can achieve its objectives in less speech restrictive ways, it must pursue those avenues instead of regulating speech based on content.⁴⁷³ However, less constitutionally suspect solutions will not protect voters and democracy from political advertisement deepfakes.

In *Alvarez*, the Justices determined that the statute was unconstitutional because there were less speech prohibitive solutions to diminish the harm from military honors lies.⁴⁷⁴ In its decision, the Supreme Court mentioned counterspeech could successfully counteract the lie.⁴⁷⁵ However, counterspeech is insufficient to offset deepfakes because it relies upon the faulty theory that true speech can counteract false speech.⁴⁷⁶ Additionally, it fails to account for the complexities that AI technology introduces into the marketplace of ideas.

The idea that the truth prevails in the marketplace of ideas is not realistic in the digital age.⁴⁷⁷ Falsehoods may at one time have been successfully diminished through counterspeech.⁴⁷⁸ However, scholars such as Professor Blitz and Professor Cass R. Sunstein argue that in modern times, the premise

466. *See supra* notes 13–16 and accompanying text.

467. Brett Gordon, Mitchell J. Lovett, Bowen Luo & James Reeder, *How Much Do Campaign Ads Matter?*, KELLOGG INSIGHT (Nov. 1, 2021), <https://insight.kellogg.northwestern.edu/article/how-much-do-campaign-ads-matter> [<https://perma.cc/A3YP-RBNX>].

468. *See supra* notes 328–33 and accompanying text.

469. *See supra* Parts I.C.2.c, II.A.4.

470. *See supra* notes 434–36 and accompanying text.

471. *See infra* Part III.C.

472. *See supra* notes 322–27 and accompanying text.

473. *See supra* note 105 and accompanying text.

474. *See supra* note 166 and accompanying text.

475. *See supra* note 166 and accompanying text.

476. *See supra* notes 379–81 and accompanying text.

477. *See* Blitz, *supra* note 231, at 223–24.

478. *See id.*

of the marketplace of ideas is no longer valid.⁴⁷⁹ First, AI technology complicates the operation of the marketplace and the effectiveness of counterspeech.⁴⁸⁰ When the verification of authenticity is difficult—if not impossible—counterspeech is not practical because voters cannot discern what speech is authentic, let alone true.⁴⁸¹ Second, with the rise of computer “bots,” false claims can appear to have widespread support when in reality they are amplified and spread through algorithms, not people.⁴⁸² Accordingly, what seems to be “true” speech, based upon widespread support, is promoted artificially and does not reflect truth.⁴⁸³ Therefore, counterspeech fails to diminish the impact that a deepfake has on a voter.

Imposing disclosure requirements is a commonly suggested less-restrictive alternative for regulating deepfakes.⁴⁸⁴ Other proposed and enacted state, congressional, and agency actions have suggested implementing disclosure requirements on AI-generated political media.⁴⁸⁵ Although disclosures present different constitutional questions regarding compelled speech,⁴⁸⁶ they also will be unsuccessful in curbing the dangers of deepfakes.

Commentators have indicated disclosures may prove futile when bad actors can edit them out and further spread misinformation.⁴⁸⁷ Moreover, flooding the marketplace with disclosures protects those who create advertisements with malicious motivations.⁴⁸⁸ For example, simple text disclosures can create the misconception that advertisements without labels are not manipulated.⁴⁸⁹ As a result, creators of deepfake advertisements intending to deceive voters or influence elections can forego disclosures and cause further harm.⁴⁹⁰ There is also no guarantee that voters will notice or understand what the disclosure intends to communicate.⁴⁹¹ Even if

479. *See id.*; *see also* Sunstein, *supra* note 22, at 393.

480. *See* Green, *supra* note 21, at 1484 n.237 (“If verification of authenticity is nearly if not entirely impossible, then counterspeech is ineffective.”).

481. *See id.*

482. *See* Blitz, *supra* note 231, at 223–24.

483. *See id.*

484. *See supra* note 325 and accompanying text.

485. *See supra* notes 47, 50, 67, 70–71, 89 and accompanying text.

486. Freedom of speech under the First Amendment applies when the government restricts speech and when it compels a person to carry an unwanted message, such as requiring advertisements to contain disclosures. *See* Abner S. Greene, “Not in My Name” *Claims of Constitutional Right*, 98 B.U. L. REV. 1475, 1486–90 (2018).

487. *See* Laura Brown, *Law Prof: Deepfakes Bill May Avoid 1st Amendment Issues*, MINN. LAW. (May 24, 2023), <https://minnlawyer.com/2023/05/24/law-prof-deepfakes-bill-may-avoid-1st-amendment-issues/> [<https://perma.cc/8GYM-RYWV>]; *see also* Weiner & Norden, *supra* note 91.

488. *AI and the Future of Our Elections*, *supra* note 141, at 19–20 (prepared statement of Ari Cohn, Counsel, TechFreedom).

489. Katerina Cizek & Shirin Anlen, *The Thorny Art of Deepfake Labeling*, WIRED (May 5, 2023, 8:00 AM), <https://www.wired.com/story/the-thorny-art-of-deepfake-labeling/> [<https://perma.cc/UW59-D5HQ>].

490. *AI and the Future of Our Elections*, *supra* note 141, at 19–20 (prepared statement of Ari Cohn, Counsel, TechFreedom).

491. Weiner & Norden, *supra* note 91.

accompanied by a disclosure, voters seeing or hearing a candidate say or do something that they disagree with may still change their behavior accordingly. Thus, disclosures are ineffective at diminishing the harms of federal candidate deepfakes in advertisements.

Similarly, requiring campaigns or parties to abide by a pledge to refrain from using AI technology will be futile. Absent any enforcement or liability, there are no incentives for these actors to comply with a pledge.⁴⁹² Also, requiring a candidate or campaign to make a pledge leaves out bad actors that have less direct ties to campaigns and ethical obligations, like PACs and super PACs.⁴⁹³ These organizations can engage in more ethically questionable tactics because they are subject to less oversight than a candidate, campaign, or party.⁴⁹⁴ Consequently, a pledge to not use AI-generated deepfakes of candidates will be ineffective.

Lastly, requiring public education in media literacy is an inadequate solution. For more rudimentary doctored media, such as “cheap fakes,” education could help the public decipher doctored media from real media. However, deepfakes have reached an unprecedented level of realism.⁴⁹⁵ Detecting whether a video is real or AI-generated has become increasingly complex, even for those highly educated in technology.⁴⁹⁶ Further, given the rapid development of AI technology, techniques to determine if a video or image is a deepfake may soon be obsolete.⁴⁹⁷ Therefore, when deepfakes are particularly harmful—such as when they depict a federal candidate—and challenging to detect, their prohibition is necessary.⁴⁹⁸

C. Recommended Revisions to the Protect Elections from Deceptive AI Act

Although the proposed ban could pass constitutional scrutiny as written, there are three revisions that would strengthen the bill’s ability to overcome a First Amendment challenge. To address the harm at issue more directly, changes should be considered to the bill’s intent and time requirements. First, the statutory requirement that an individual acts with the purpose of “influencing an election or soliciting funds” should be more specific to hold bad actors accountable.⁴⁹⁹ Second, the ban should be limited to a fixed

492. See Manzi, *supra* note 269, at 2642.

493. See Tim Lau, *Citizens United Explained*, BRENNAN CTR. FOR JUST. (Dec. 12, 2019), <https://www.brennancenter.org/our-work/research-reports/citizens-united-explained> [<https://perma.cc/4RLP-MZVK>] (explaining how the impact of the *Citizens United v. Federal Election Commission*, 558 U.S. 310 (2010), decision permitted PACs and super PACs to circumvent regulations).

494. See *id.*

495. See Chesney & Citron, *supra* note 12, at 1759.

496. See Drew Harwell, *Top AI Researchers Race to Detect ‘Deepfake’ Videos: ‘We Are Outgunned’*, WASH. POST (June 12, 2019, 4:44 PM), <https://www.washingtonpost.com/technology/2019/06/12/top-ai-researchers-race-detect-deepfake-videos-we-are-outgunned/> [<https://perma.cc/NNM3-326K>].

497. See *id.*

498. See *supra* notes 426–27 and accompanying text.

499. See Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

period before the election to narrow its application and focus on the most damaging advertisements.⁵⁰⁰ Lastly, enforcement by an independent agency, such as the FEC, should be considered to mitigate concerns of partisan enforcement.

It is difficult to ascertain whether someone had the purpose of “influencing an election.”⁵⁰¹ The intent standard should instead reflect what kind of bad actors the statute seeks to target. The bill should incorporate the intent requirements from the recently enacted Michigan election deepfake statute.⁵⁰² The Michigan law requires an actor to intend that the deepfake’s distribution will (1) “harm the reputation or electoral prospects of a candidate in an election” and (2) change the voting behavior of electors “into incorrectly believing that the depicted individual engaged in the speech or conduct depicted.”⁵⁰³ Requiring an actor to disseminate a deepfake for the purpose of “influencing an election” is too vague and difficult to establish. Instead, the proposed ban should emphasize intended voter deceit to tailor liability to a more specific harm.

Second, the Protect Elections from Deceptive AI Act does not contain a period when it is applicable.⁵⁰⁴ Adding a time limitation to narrow the legislation is imperative. In *Citizens United v. Federal Election Commission*,⁵⁰⁵ the Supreme Court upheld disclosure requirements for independent expenditures and electioneering communications.⁵⁰⁶ This included permitting limits based on the timing of election speech in the campaign finance context.⁵⁰⁷ In its decision, the Court allowed regulation of “electioneering communications” within thirty days of a primary election and sixty days before a general election.⁵⁰⁸ Deepfake advertisements cause the most harm when there is insufficient time for the deepfake to be taken down and labeled as false.⁵⁰⁹ Thus, limiting the prohibition to a fixed period before the election, as much of the proposed federal and state legislation does, would narrow the ban’s application and make it more likely to pass scrutiny.⁵¹⁰ Adopting the time limitations of other proposed and enacted legislation, the

500. See Weiner & Norden, *supra* note 91 (stating broader prohibitions of electoral deepfakes should be restricted to preelection windows).

501. See S. 2770, § 2.

502. See *supra* note 47 and accompanying text.

503. See *supra* note 47 and accompanying text.

504. See S. 2770.

505. 558 U.S. 310 (2010).

506. The decision invalidated the Bipartisan Campaign Reform Act (BCRA) provision that prohibited corporations from using their general treasury funds for express advocacy or electioneering communications. *Id.* at 365–66. However, the Court upheld the BCRA’s provision regarding disclaimer and disclosure requirements because they were a “less restrictive alternative to more comprehensive regulations of speech.” *Id.* at 369.

507. *Id.* at 367–71.

508. *Id.* at 337, 367–71.

509. Jack Langa, *Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes*, 101 B.U. L. REV. 761, 789 (2021) (“[A] temporal limit allows for greater creative expression through deepfake technology but narrowly targets instances when the deepfakes could do the most harm.”).

510. See *supra* notes 48–49, 47, 58, 73 and accompanying text.

Protect Elections from Deceptive AI Act should apply sixty days before a primary election or ninety days before a general election. This will ensure that the bill targets the most harmful deepfake advertisements and does not impermissibly chill protected speech.

Third, the negative theory arguments regarding fears of partisan prosecution are significant. The bill permits depicted candidates to obtain injunctive or monetary relief against deepfake creators and distributors in federal court.⁵¹¹ Instead of enforcement by government actors politically accountable to the President, the proposed deepfake advertisement ban could be implemented by an independent, non-partisan government agency.⁵¹²

The FEC is a federal agency that could have jurisdiction regarding deepfakes in the context of elections.⁵¹³ However, the FEC's path to regulation is not without hurdles.⁵¹⁴ First, it currently has limited jurisdiction, which narrows its ability to regulate in this space.⁵¹⁵ Its central focus in regulatory efforts pertains to increasing transparency regarding sponsorship and funding for political advertising through disclosure.⁵¹⁶ Second, scholars raise concerns regarding the FEC's dysfunction and ineffectiveness at holding bad actors accountable.⁵¹⁷ Nevertheless, enforcement through the FEC is a worthwhile consideration, given fears of partisanship when regulating false political speech.

Scholars evaluating deepfake regulation have suggested implementing monetary penalties or some right of retraction through an independent government agency to limit political deepfakes more neutrally.⁵¹⁸ Recent regulatory efforts have sought to expand the FEC's jurisdiction to regulate deepfakes in political advertisements.⁵¹⁹ However, the proposed FEC action permits the deepfake creator to use disclosure to avoid liability. As mentioned, disclosures are not an adequate solution to deepfakes of federal candidates in advertisements.⁵²⁰ Because prohibiting these harmful deepfakes is necessary, attempts to lessen adverse consequences, such as politicized enforcement, should be pursued. Therefore, legislators should consider whether the FEC should implement the proposed ban to lessen concerns of partisan abuse.

511. See Protect Elections from Deceptive AI Act, S. 2770, 118th Cong. § 2 (2024).

512. Sunstein, *supra* note 22, at 418.

513. Chesney & Citron, *supra* note 12, at 1807–08.

514. See *id.*

515. See *id.* at 1807.

516. For the FEC to regulate deepfakes, an amendment may be required to broaden the FEC's jurisdiction under 52 U.S.C. § 30106(b)(1). See O'Donnell, *supra* note 53, at 724 n.226.

517. See Adam Skaggs, *A Call to Abolish the FEC*, BRENNAN CTR. FOR JUST. (Dec. 2, 2011), <https://www.brennancenter.org/our-work/research-reports/call-abolish-fec> [<https://perma.cc/2W36-33CT>].

518. See Sunstein, *supra* note 22, at 418.

519. See *supra* Part I.B.2.

520. See *supra* Part III.B.

CONCLUSION

Deepfakes of federal candidates in political advertisements pose a significant threat to voters and democracy. Despite competing First Amendment protections for false speech and political speech, the Protect Elections from Deceptive AI Act is constitutional. The proposed ban fits into the exception for regulating false speech that causes a “legally cognizable harm” to identifiable victims in *Alvarez*. Also, despite implicating political discourse, there is a compelling government interest in limiting deepfakes to safeguard elections. AI-generated deepfakes of federal candidates present unique challenges that warrant their prohibition in advertisements. Moreover, the bill is narrowly tailored and is necessary to address the harm deepfakes pose. Thus, the Protect Elections from Deceptive AI Act is constitutional because it offers a narrowly crafted solution to a novel harm, and successfully balances the risk to democracy against freedom of speech. The recommended revisions and proposals more precisely tailor the statute to withstand strict scrutiny and lessen concerns of partisan enforcement.