Concealed Carry and the Right to Bear Arms
— Joseph Greenlee

The Fifth Amendment’s Act of Production Doctrine: An Overlooked Shield Against Grand Jury Subpoenas Duces Tecum
— Peter Thomson

The Future of the Federal Judiciary
— Diarmuid F. O'Scanlайн

From Orange Groves to Cryptocurrency: How Will the SEC Apply Longstanding Tests to New Technologies?
— Troy Paredes, Scott Kimpel

Website Inaccessibility: The New Wave of ADA Title III Litigation
— J. Gregory Grisham

The Supreme Court Takes Up Abortion: What You Need to Know About June Medical Services v. Gee
— Rachel N. Morrison

Is It Time to Revisit the Constitutionality of Unauthorized Practice of Law Rules?
— Michael E. Rosman

The Ministerial Exception After Hosanna-Tabor: Firmly Founded, Increasingly Refined
— J. Gregory Grisham, Daniel Blomberg
Letter from the Editor

The **Federalist Society Review** is the legal journal produced by the Federalist Society for Law & Public Policy Studies. The Review features excellent scholarship on important legal and public policy issues from some of the best legal minds in the country.

**Review** articles are published thanks to the hard work of our fifteen Practice Group Executive Committees and authors who volunteer their time and expertise. The Review seeks to contribute to the marketplace of ideas in a way that is collegial, accessible, intelligent, and original. Articles are available at [fedsoc.org](http://fedsoc.org) and through the Westlaw database.

We hope that readers enjoy the articles and come away with new information and fresh insights. Please send us any suggestions and responses at info@fedsoc.org.

Sincerely,

Katie McClendon
Director of Publications
The Federalist Society Review

Volume 20  2019

The Fifth Amendment’s Act of Production Doctrine: An Overlooked Shield Against Grand Jury Subpoenas

Duces Tecum

by Peter Thomson ................................................................. 4

Can and Should the Federal Judiciary Rein In Our Expansive Administrative State?

by Ted Hirt ........................................................................... 10

Lessons in Reading Law: *Rimini Street’s* Duel Over “Full Costs”

by Joel S. Nolette .................................................................. 18

John Marshall’s Jurisprudence Supports Preemption of California’s Net Neutrality Law

by Randolph J. May, Seth L. Cooper ...................................... 26

Concealed Carry and the Right to Bear Arms

by Joseph Greenlee ............................................................... 32

Protecting Individual Liberty Through State Constitutional Law: Judge Sutton’s Plea for Federalism in Judicial Decisionmaking

by John C. O’Quinn, Jason M. Wilcox .................................... 38

The Tenacity of Transformation Theory, and Why Constitutional History Deserves Better

by Stephen B. Presser ........................................................... 42

Can a New Establishment Clause Jurisprudence Succeed in Protecting Religious Minorities Where *Lemon* Has Failed?

by Alexandra M. McPhee ...................................................... 46

From Orange Groves to Cryptocurrency: How Will the SEC Apply Longstanding Tests to New Technologies?

by Troy Paredes, Scott Kimpel .............................................. 56

Website Inaccessibility: The New Wave of ADA Title III Litigation

by J. Gregory Grisham .......................................................... 66

Is It Time to Revisit the Constitutionality of Unauthorized Practice of Law Rules?

by Michael E. Roseman ....................................................... 74

The Ministerial Exception After *Hosanna-Tabor*: Firmly Founded, Increasingly Refined

by J. Gregory Grisham, Daniel Blumberg .............................. 80

More News on Powers Reserved Exclusively to the States

by Robert G. Natelson .......................................................... 92

Whistling in *Chevronland*: Why Department of Labor Interpretations of the Sarbanes-Oxley Act Whistleblower Provisions Do Not Deserve Judicial Deference

by Donn C. Meindertsma ...................................................... 100

Have the American People Irrevocably Ceded Control of Their Government to the Modern Administrative State?

by Ted Hirt ........................................................................... 112

Two Views on Criminal Justice Reform: The Author and a Critic on *Locked In*

by Vikrant P. Reddy, Kent Scheidegger .................................. 120

The Future of the Federal Judiciary

by Diarmuid F. O’Scaunnla ................................................... 138

The Supreme Court Takes Up Abortion: What You Need to Know About *June Medical Services v. Gee*

by Rachel N. Morrison ......................................................... 144

*Carpenter v. United States*: A Reevaluation of First Principles, One Year On

by Dean A. Mazzone ............................................................ 150

Civility, Courage, Humility: Lessons In Life and Law From a Justice

by Elizabeth Slattery ............................................................ 158

National Security, Trade, and China: Risks and Responses in the Internet Age

by Ronald A. Cass ............................................................... 162

An Unconstitutional Attempt to Address Affordable Housing

by Jeffrey M. Harris ............................................................. 168

Credentials Not Required: Why an Employee’s Significant Religious Functions Should Suffice to Trigger the Ministerial Exception

by Nathaniel M. Fouch, Erik Money, & Thomas C. Berg ....... 182

*Kisor v. Wilkie* Makes *Auer* a Paper Tiger

by Karen Harned, Aeron Von Scoly ....................................... 192
The Fifth Amendment’s Act of Production Doctrine: An Overlooked Shield Against Grand Jury Subpoenas Duces Tecum

By Peter Thomson

Criminal Law & Procedure Practice Group

About the Author:
Peter Thomson is a Member of the law firm Stone Pigman Walther Wittmann LLC in New Orleans. He previously served for 23 years as a federal prosecutor with the U.S. Department of Justice.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. We invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

When asked whether a company can assert the Fifth Amendment and refuse to produce documents demanded by grand jury subpoena, most criminal defense attorneys would answer, “No.” And they would be right: a company—any legal entity such as a corporation, partnership, or L.L.C. (collectively “company”)—has no privilege against self-incrimination under the Fifth Amendment, regardless of whether the contents of the subpoenaed documents incriminate the company.1 Furthermore, except in some cases involving sole proprietorships,2 an individual who produces documents on behalf of a company generally also has no Fifth Amendment protection, even where the contents of subpoenaed documents incriminate the individual personally. In either situation, the corporate representative must produce the incriminating documents to the government.

Similarly, when asked whether an individual has a Fifth Amendment right to refuse to produce private documents demanded by grand jury subpoena, issued to the individual in his personal capacity, many criminal defense attorneys also might answer, “No.” But they would not be entirely correct. Although the Fifth Amendment generally does not shield the incriminating contents of private documents, criminal practitioners often overlook, and sometimes misunderstand, a somewhat elusive jurisprudential rule called the “act of production doctrine.” Under the doctrine, an individual can assert his Fifth Amendment privilege against self-incrimination and refuse to produce subpoenaed documents where the act of producing them is incriminating in itself, regardless of the contents of the documents. The doctrine is based on the concept that, in certain situations, the very act of disclosing documents to the government can have a testimonial aspect which, if compelled and incriminating, is equivalent to compelled incriminating oral testimony, which is protected by the Fifth Amendment.

I. The Fifth Amendment

The government violates the Fifth Amendment when it seeks to compel an individual to testify to information that can be used to prosecute the individual for a crime or which provides a link in the chain of evidence needed to prosecute. The Fifth Amendment provides, in pertinent part, that “[n]o person . . . shall be compelled in any criminal case to be a witness against himself.”3 The privilege against self-incrimination extends not only to answers that would in themselves support a conviction under a federal criminal statute but likewise embraces those which would

---

2 See Braswell, 487 U.S. at 118 n.11 (1988) (leaving open the question of whether the agency rationale supports compelling a custodian to produce documents where the custodian establishes that he is the sole officer and employee of the entity and the jury would inevitably conclude he produced the records).
3 U.S. Const. amend. 5.
furnish a link in the chain of evidence needed to prosecute . . . [an individual] for a federal crime."4 The mere possibility of criminal prosecution is all that is needed to properly invoke the privilege.5 In fact, a witness may properly assert the Fifth Amendment while simultaneously maintaining his innocence. This is true because the privilege protects even "innocent men . . . who otherwise might be ensnared by ambiguous circumstances."6

In order to successfully assert the privilege against self-incrimination, a witness must demonstrate that the information sought by the government is 1) compelled, 2) incriminating, and 3) testimonial.7 The first two elements—compelled and incriminating—are rarely at issue. "Compelled" simply means not voluntarily given. "Incriminating" means that the information demanded tends to show guilt or furnishes a link in a chain of evidence needed to prosecute. However, the meaning of the third element—testimonial—is not so clear and has been the focus of much debate by scholars and in the courts.

Although the Supreme Court has yet to establish a bright line rule for determining when a witness’ statement is "testimonial," the Court nevertheless has provided important guidance by holding that a statement is testimonial when it relates to an assertion of fact.8 Thus, in order to qualify as testimonial, a witness’ "communication must itself, explicitly or implicitly, relate a factual assertion or disclose information. Only then is a person compelled to use the contents of his own mind to assert a fact."9 Even though these acts communicate information, none are considered to be "testimonial" within the meaning of the Fifth Amendment. None require a person to use the contents of his own mind to assert a fact.

With respect to government demands for documents, the Supreme Court, in an early landmark case addressing privacy rights, held that the Fifth Amendment’s privilege against self-incrimination protects against the compelled production of any incriminating documents.10 However, twenty years later, in 1906, the Supreme Court declined to extend the privilege to corporations responding to grand jury subpoenas.11 It is now settled law that the Fifth Amendment does not protect the contents of business records,12 which are, for the most part, voluntarily prepared documents and therefore not compelled. Furthermore, under what is known as the "collective entity doctrine," a company has no Fifth Amendment privilege against self-incrimination. Except in some cases of sole proprietorships, which do not exist independently of the persons who comprise them,13 the right to resist compelled self-incrimination is "a 'personal privilege,' which companies and other collective entities do not share."14 This is true regardless of whether a document produced incriminates the company or its records custodian.15 Thus, a custodian who produces records on behalf of a company "may not resist a subpoena for corporate records on Fifth Amendment grounds."16

Likewise, the contents of privately held documents are not protected by the Fifth Amendment, unless the government compels their creation or requires the witness to endorse the truth of their incriminating contents.17 Otherwise, the government may compel the production of private papers.18 For example, if a document was voluntarily prepared prior to the issuance of a subpoena, it must be produced in response to the subpoena because it is not a compelled statement within the meaning of the Fifth Amendment. Previously created personal calendars, appointment books, day planners, personal journals, diaries, or other similar personal documents or papers, as well as documents prepared to comply with state or federal regulations, such as tax returns, are all examples of private documents and papers the contents of which generally are unprotected by the Fifth Amendment.19 As the Supreme Court has summarized, "the Fifth Amendment protects against 'compelled self-incrimination, not (the disclosure of) private information.'"20

However, a grand jury’s subpoena power is not unlimited and may not “violate a valid privilege, whether established by the Constitution, statutes, or the common law.”21 In evaluating whether a grand jury subpoena might violate the Fifth Amendment, courts often examine the objects and scope of the

---

5 In re Seper, 705 F.2d 1499, 1501 (9th Cir. 1983).
9 Id. at 210.
10 Id.
11 Id.
demand as well as the method of production requested by the government.

II. The Act of Production Doctrine

Although an individual cannot assert the Fifth Amendment's privilege against self-incrimination to shield the contents of pre-existing, voluntarily created documents, the act of production doctrine recognizes that the Fifth Amendment protects an individual from being compelled to produce documents (i.e., any written materials, including emails and text messages) in response to a subpoena where the act of production itself implicitly has a testimonial aspect. Considering the facts of a given case, the compelled production of documents may communicate "statements of fact" that incriminate the person producing them, including that the documents (1) exist, (2) are in the person's possession or control, and (3) are authentic. Thus, by merely disclosing subpoenaed documents to the government, a witness might effectively be "testifying" to factual information that could be used by the government against that witness, either directly or through the development of investigative leads. In determining whether an act of production is testimonial, federal courts tend to focus on whether the existence of the documents at issue was known to the government at the time of the subpoena's issuance, in which case the witness would merely be "surrendering" them as opposed to testifying to their existence, location, or authenticity. The act of production doctrine is not available to companies because the Fifth Amendment privilege is a personal one.

A. The Doctrine's Development

The act of production doctrine derives principally from the Supreme Court's 1976 decision in Fisher v. United States. In that case, Fisher asserted the privilege against self-incrimination after the Internal Revenue Service (IRS) served a summons on his lawyer for certain tax records prepared by his accountant. The Supreme Court held that the contents of the subpoenaed records, though possibly incriminating, were not protected because they had been voluntarily prepared before the subpoena was issued and thus were not "compelled" within the meaning of the Fifth Amendment. Hence, Fisher could not prevent the records from being produced solely because they contained incriminating evidence against him, regardless of whether the records belonged to him or someone else.

However, the Court recognized that the Fifth Amendment is implicated when the act of complying with a subpoena is both "testimonial" and "incriminating." The Court explained that "[t]he act of producing evidence in response to a subpoena nevertheless has communicative aspects of its own, wholly aside from the contents of the papers produced. Compliance with the subpoena tacitly concedes the existence of the papers demanded and their possession or control by the taxpayer." Although Fisher implicitly admitted to the existence and possession of the records by disclosing them, the Court concluded that the disclosure was not "testimonial" because the existence and location of the records sought by the government was a "foregone conclusion." Thus, because the IRS already knew of the existence of the records, and where to find them, their disclosure did not communicate any new information that incriminated Fisher. According to the Court, "the taxpayer add[ed] little or nothing to the sum total of the Government's information" by conceding that he in fact had the records at issue. Compliance with the summons therefore was a question "not of testimony but of surrender." Fisher, accordingly, instructs us to ask 1) whether the act of production communicates the existence, control, or authenticity of the document produced and 2) whether the incriminating factual information communicated provides the government with evidence it might otherwise not have. If both elements are met, then the act of production is testimonial.

In a subsequent landmark case, United States v. Doe, the Supreme Court held that Doe validly invoked the privilege against self-incrimination in refusing to produce documents subpoenaed by a federal grand jury because his compliance "would involve testimonial self-incrimination." The subpoenas, which were drafted in sweeping terms, demanded the production of business records of a sole proprietorship through which Doe conducted business. The trial court held that Doe's compliance with the subpoenas would infringe on his Fifth Amendment rights because it would require him to admit that the records existed, were in his possession, and were authentic. The Court of Appeals agreed, finding no proof in the record that the government knew that the records were in Doe's possession or control prior to issuing the subpoenas. In fact, the Court accused the government of trying to compensate for its lack of information by demanding that Doe be an informant against himself. The Supreme Court agreed, holding that Doe's act of compliance would necessarily involve testimonial self-incrimination, against which he was protected by the Fifth Amendment.

In United States v. Hubble, the Supreme Court expounded on Fisher's "foregone conclusion" analysis in the context of the infamous Clinton Whitewater investigation. Hubble was served with a subpoena demanding the production of a vast number of

24 Hubbell, 530 U.S. at 36.
25 Id.
26 425 U.S. at 408.
27 Id. at 409-10.
28 Id.
29 Id. at 410.

30 Id.
31 Id. at 411.
32 Id.
33 Id. (quoting In re Harris, 221 U.S. 274, 279 (1911)).
34 465 U.S. 605.
35 Id. at 613.
36 Id.
37 Id.
38 Id.
39 530 U.S. 27.
documents spanning a several-year time period. After he asserted the Fifth Amendment privilege and refused to comply with the subpoena, the government granted him immunity, obtained the documents, and then indicted him based on the contents of the documents. The Supreme Court held that Hubbell could not be prosecuted based on the contents of the documents because the government had made derivative use of the testimony implied by their production during its investigation that led up to the criminal charges. Thus, the government was unable to demonstrate that the evidence it used to obtain the indictment was "wholly independent" of (i.e., not derivatively sourced from) Hubbell's immunized testimonial act of subpoena compliance.

In dismissing the indictment, the Court reasoned that the existence and locations of the documents sought by the government were not a "foregone conclusion" at the time the subpoena was issued to Hubbell. The Court stressed that the government cannot prosecute individuals based on incriminating materials obtained through "fishing expeditions" conducted using grand jury subpoenas duces tecum—subpoenas for the production of documents spanning a several-year time period. After he asserted the Fifth Amendment privilege and refused to comply with the subpoena, the government granted him immunity, obtained the documents, and then indicted him based on the contents of the documents because the government had made derivative use of the testimony implied by their production during its investigation that led up to the criminal charges. Thus, the government was unable to demonstrate that the evidence it used to obtain the indictment was "wholly independent" of (i.e., not derivatively sourced from) Hubbell's immunized testimonial act of subpoena compliance.

In dismissing the indictment, the Court reasoned that the existence and locations of the documents sought by the government were not a "foregone conclusion" at the time the subpoena was issued to Hubbell. The Court stressed that the government cannot prosecute individuals based on incriminating materials obtained through "fishing expeditions" conducted using grand jury subpoenas duces tecum—subpoenas for the production of documents spanning a several-year time period. After he asserted the Fifth Amendment privilege and refused to comply with the subpoena, the government granted him immunity, obtained the documents, and then indicted him based on the contents of the documents because the government had made derivative use of the testimony implied by their production during its investigation that led up to the criminal charges. Thus, the government was unable to demonstrate that the evidence it used to obtain the indictment was "wholly independent" of (i.e., not derivatively sourced from) Hubbell's immunized testimonial act of subpoena compliance.

In dismissing the indictment, the Court reasoned that the existence and locations of the documents sought by the government were not a "foregone conclusion" at the time the subpoena was issued to Hubbell. The Court stressed that the government cannot prosecute individuals based on incriminating materials obtained through "fishing expeditions" conducted using grand jury subpoenas duces tecum—subpoenas for the production of documents spanning a several-year time period. After he asserted the Fifth Amendment privilege and refused to comply with the subpoena, the government granted him immunity, obtained the documents, and then indicted him based on the contents of the documents because the government had made derivative use of the testimony implied by their production during its investigation that led up to the criminal charges. Thus, the government was unable to demonstrate that the evidence it used to obtain the indictment was "wholly independent" of (i.e., not derivatively sourced from) Hubbell's immunized testimonial act of subpoena compliance.

B. Raising the Doctrine as a Shield in Practice

Under the right circumstances, the act of production doctrine can be a formidable shield against government compelled disclosure of private documents where the production would lead to the government's discovery of the existence of documents, the subpoenaed party's possession of them, or the belief by the witness that the documents are responsive (i.e., authentic). Nonetheless, when faced with subpoenas duces tecum, many criminal defense attorneys do not consider—much less raise—the act of production doctrine. Most motions to quash subpoenas are based on the grounds of vagueness, overbreadth, unreasonableness, and/or undue burden on the responding witness. As a practical matter, however, most of those challenges have limited success, largely due to the wide latitude given to federal prosecutors and grand juries by the courts.

There are several possible reasons criminal practitioners might neglect to invoke the act of production doctrine as a shield against grand jury subpoenas. First, it might be due to a misunderstanding of the Supreme Court's decision in Fisher, which declined to extend Fifth Amendment protection to private tax records held by an individual. Second, it might be due to a misunderstanding of how the mere production of documents can be protected under the Fifth Amendment when the contents of the same documents are not. Distinguishing the act of producing the records from their incriminating contents as the basis of the Fifth Amendment claim can be difficult, particularly since invoking the act of production doctrine with respect to documents also may effectively shield their incriminating contents. Third, in situations where an act of production would provide the government with a link in the chain of evidence needed to prosecute the witness, an even more complicated picture can arise. Because the doctrine is grounded in the Fifth Amendment, it protects not only acts of production which are intrinsically incriminating to the subpoenaed witness, but also productions which provide the government with a link in the chain of evidence needed to prosecute. Finally, because of the lack of judicial clarity regarding the meaning of "testimonial," criminal practitioners might find it difficult to discern the circumstances under which an act of production is protected by the Fifth Amendment, particularly since no bright line test has been established by the Supreme Court.

Notwithstanding the lingering ambiguity, since Hubble it appears that the determination of whether an act of production qualifies as a "testimonial communication" turns on the foregone conclusion test, which asks whether the government knew of the existence and location of the subpoenaed documents prior to their production. Categorical requests for documents the government believes are likely to exist are not sufficient. Although the government apparently does not have to show "actual knowledge" of the existence of each and every document described, it nevertheless must establish its knowledge with "reasonable particularity," not merely infer that the documents exist and are in the control of the subpoenaed party. Only where the government can make such a showing is the production not considered to be compelled testimony protected by the Fifth Amendment.

Nonetheless, grand jury subpoenas sometimes are drafted so broadly that they encompass almost all conceivable written materials under a person's control. For example, a subpoena might demand that an individual produce any and all emails and texts in his possession, or call for the production of "any and all documents" related to a certain entity or subject matter where the term "document" is defined in excessively broad if not almost limitless terms. But such expansively worded subpoenas are classic examples of the type of fishing expeditions that have been

---

40 Id. at 43.
41 Id.
struck down time and again as unconstitutional. Complying with such an exhaustive list of demands, in some cases, can be “tantamount to answering a series of interrogatories asking a witness to disclose the existence and location of particular documents fitting certain broad descriptions,” which was condemned by the Court in *Hubbell*. Further, the fact that a demand might be limited to documents or records related to a specific entity or subject matter does not necessarily cure its unconstitutionality where a response might reveal the existence, location, or authenticity of documents unknown to the government.

Moreover, the more broadly a subpoena is written, the more likely the government is unable to identify specific information that both exists and is in the individual’s possession. In such instances, the government might not merely be asking for the surrender of documents actually known to it, but instead might be employing the subpoenaed party to help unearth additional evidence against that party. The government cannot subpoena the “testimonial aspect” of a person’s production of information to use it as a road map to uncover evidence which can be used against him.

Finally, grand jury subpoenas also sometimes actually demand that witnesses generate certain types of documents, including compilations. A subpoena might demand a list of information predicated, or not, on the contents of preexisting written materials in the witness’ possession. These could be lists of all bank accounts, certain items, relationships with other persons, interests in certain investments, or names of legal entities the witness has an interest in. In all of these examples, the act of producing the subpoenaed information would be testimonial in character and, if potentially incriminating, protected under the Fifth Amendment.

III. Conclusion

In order to successfully assert the Fifth Amendment, a subpoenaed party must show that the information sought by the government is compelled, incriminating, and testimonial in character. The meaning of “testimonial” often is at issue, particularly in the context of producing documents demanded by a subpoena duces tecum. In determining whether an act of producing documents has a testimonial aspect protected by the Fifth Amendment, courts now seem to focus on the foregone conclusion test, which prohibits categorical demands for documents the government believes but does not know exist. The pivotal question under the foregone conclusion test is whether the existence and location of the documents at issue were known to the government with reasonable particularity at the time the subpoena issued. If not, then the witness’ act of producing the documents could communicate information that the government did not have, such as the existence, possession, and authenticity of the subpoenaed documents. The government cannot try to use the witness’ testimony—inherent in the act of complying with the subpoena—against him in violation of the Fifth Amendment. In such instances, the act of production doctrine can be raised as a shield against government compelled disclosure of private documents.


50 *Doe*, 465 U.S. at 614 n.12.

51 See *Hubbell*, 530 U.S. at 42.
Can and Should the Federal Judiciary Rein In Our Expansive Administrative State?

By Ted Hirt

Administrative Law & Regulation Practice Group

A Review of:

Judicial Fortitude: The Last Chance to Rein In the Administrative State, by Peter J. Wallison

About the Author:

Ted Hirt is a Professorial Lecturer in Law at the George Washington University Law School. He previously was a career attorney in the United States Department of Justice’s Civil Division for almost 37 years. The views he expresses are his own.

Note from the Editor:

The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:


Conservative legal scholars have been engaged in rethinking the modern administrative state, including how its expansive reach might be constrained or reversed.¹ The most recent contribution to this intellectual ferment is Peter J. Wallison’s provocative book, Judicial Fortitude: The Last Chance to Rein In the Administrative State. As its title suggests, this book advocates a recalibration of our Constitution’s structural separation of powers under which the federal judiciary would constrain the excesses of administrative agency powers.² Wallison concludes that Congress has been derelict in its responsibility to perform this critical task.³ He recommends that the judiciary act to protect the reservation of exclusive lawmaking responsibility to Congress by prohibiting agency efforts to resolve legislative issues.⁴ Otherwise, he fears that there will be more power “concentrated in a faceless bureaucracy” and a resulting loss of public confidence in the democratic processes that govern our society.⁵

Wallison is a Senior Fellow at the American Enterprise Institute in Washington, D.C. In addition to having practiced law in Washington, D.C. and New York City, Wallison brings to this book an extensive resume of government experience. From 1974 to 1976, he was counsel to then-Vice President Nelson Rockefeller; from 1981 to 1985, he was General Counsel of the Department of the Treasury; and from 1986 to 1987, he was White House Counsel for President Ronald Reagan.

Wallison’s concern is two-fold. First, he contends that Congress has failed to place sufficient limits on the exercise of powers it delegates to administrative agencies, with the result that unaccountable agencies enact and enforce rules that profoundly impact our society and national economy.⁶ Second, Wallison asserts that the judiciary has “largely surrendered its constitutional duty to determine the scope of administrative discretion.”⁷ As a consequence of these developments, he says, “we risk losing our democracy unless we can gain control of the agencies of the

³ Wallison at xiv-xvi, 39-52.
⁴ Id. at 147, 149-50.
⁵ Id. at xix, 165-66.
⁶ Id. at ix, 50-54.
⁷ Id. at ix.
administrative state.” 8 In Judicial Fortitude, Wallison develops these themes.

I. The Framers’ Vision of Limited Government Under the Rule of Law, and How That Vision Has Been Weakened and Compromised

The starting point for Wallison’s book is his judgment that the Framers’ original design of a national government of limited powers, acting within the confines of the rule of law, has been weakened by the modern administrative state. 9 The Framers of the Constitution had the “paramount goal” of preserving liberty for the American people, and crafting a government of limited powers was central to that objective. 10 A strict separation of powers was intended to preclude the tyranny that James Madison said would be effected by an “accumulation of all powers, legislative, executive and judiciary, in the same hands.” 11 That structure also enables each of the three branches to check the powers of the others. 12

According to Wallison, the nation’s founders intended that the courts would function as guardians of the Constitution’s structure and therefore of the people’s liberties; in the context of administrative agency interpretation of congressional language, he says, the courts need to guard against unwarranted expansions of agency power. 13 For Wallison, agency expansions of their own authority under the guise of statutory interpretation undermine the separation of powers and “seriously impair the rule of law.” 14 Madison, he notes, warned against the dangers of an “inconstant government,” under which the citizen (whether a “prudent merchant,” or a farmer or manufacturer) could not plan or invest in the face of arbitrary government power. 15 Under the circumstances, Wallison argues, the courts must exercise their authority “to determine and declare when either of the other branches steps outside its assigned role under the constitutional separation of powers.” 16

Agencies, Wallison argues, should not move “beyond the task of administering or enforcing the law into the role of making law, reserved by the Constitution to Congress,” but a number of agencies have done so. 17 Wallison devotes an entire chapter to what he characterizes as examples of agency overreach. 18 For example, in 1996, the Food and Drug Administration (FDA) claimed that tobacco was a “drug” that fell under its jurisdiction to regulate; this position was a reversal of a previous FDA position and was inconsistent with legislation Congress had enacted. 19 The Supreme Court held that the FDA lacked the authority to declare tobacco a drug and therefore regulate it. 20 Similarly, in 2014, in Utility Air Regulatory Group v. Environmental Protection Agency, the Supreme Court determined that the EPA had engaged in an impermissible interpretation of the Clean Air Act in its regulation of carbon dioxide emissions. 21 Wallison also cites efforts by Richard Cordray, the Director of the Consumer Financial Protection Bureau during the Obama Administration, to enforce the Dodd-Frank Act’s prohibition of abusive acts and practices in consumer finance activities by bringing enforcement actions rather than by undertaking rulemaking proceedings. 22 Wallison characterizes the agency’s reliance on enforcement actions without an underlying set of rules as a “derogation of the rule of law,” and he observes that a regulated firm “can have no idea what activities might constitute abuse and thus no way to modify its behavior.” 23 Wallison provides several other examples of what he considers to be agency overreach. 24

In a separate chapter, Wallison traces much of the expansion of agency authority in the twentieth century to the ideology of the Progressive movement of the late nineteenth and early twentieth centuries, which developed and advocated views on the Constitution and limited government that were sharply different from those of prior generations. 25 For example, in 1887, Woodrow Wilson, then an academic, rejected the separation of powers principle as inefficient, even calling it a “radical defect.” 26 For the Progressives, the rapid changes in the nation’s economy and society after the Civil War, marked by dramatic increases in corporate power and urban poverty, demanded aggressive government action outside existing constitutional structures. 27 Wallison contends that the Progressives were mistaken in concluding that our constitutional system had to be modified to accommodate the rapid societal changes, but their faith in public administration by unelected, disinterested experts nevertheless prevailed. 28 With the onset of the Great Depression, President Franklin D. Roosevelt and his New Deal allies carried forward the Progressives’ political ideas to expand the reach of the administrative state to a wide

---

8 Id.
9 Id. at xviii-xxii, 19-21.
10 Id. at 23.
11 Id. (citing The Federalist No. 47 (James Madison)).
12 Id.
13 Id. at 28-29.
14 Id. at 34.
15 Id. (citing The Federalist No. 62 (James Madison)).
16 Id. at 24-25.
17 Id. at 2.
18 Id. at 2-19.
19 Id. at 17.
22 Wallison at 2-4.
23 Id. at 3.
24 Id. at 4-9 (criticizing the Department of Education’s reliance on issuing guidance and interpretative letters in its oversight of Title IX of the Education Amendments of 1972 as evading the rulemaking process); id. at 9-11 (criticizing agencies’ use of directives to get banks to cease making loans to payday lenders).
25 Id. at 55-75.
26 Id. at 56-57 (citing Woodrow Wilson, Congressional Government in the United States 284 (1911)).
27 Id. at 57-59.
28 Id. at 58-61, 62-63.
variety of economic and social problems. 29 Although some Progressive thinkers were hesitant about letting agencies control economic regulation—especially given that some regulated entities were able to secure protections from agencies and unduly influence them, which Wallison calls “clientism” and others call “regulatory capture”—the executive agencies nevertheless succeeded in regulating many industries. 30 Wallison also observes that the Supreme Court gradually embraced Progressive views on the central role of administrative agencies in governing private conduct largely free of court interference. 31

Judicial Fortitude also questions whether the modern administrative state has yielded the benefits that its supporters have touted. 32 Wallison asserts that the American economy “is saddled with a huge number of unnecessary regulatory and administrative restraints” that are legacies of the Progressive movement and the New Deal. 33 He also says that most of our economy functions “quite well” without significant government regulation, and that the regulated part of the economy “often functions very poorly in comparison with the parts that are not regulated.” 34 Wallison contends, for example, that the nation’s banking system has been overregulated, and he criticizes the Dodd-Frank Act as creating a new round of burdensome regulations on the financial system, resulting in a significant drag on economic growth. 35 He also points to the successful deregulation of various industries (such as airlines and railroads) as evidence of the superiority of the market over burdensome regulation. 36 Wallison also cites various statistical studies in support of his contention that regulatory costs impair economic growth. An April 2016 study determined that the cumulative cost of regulations between 1980 and 2012 seriously dampened economic growth (by 25 percent); this meant that each person in the United States was nearly $13,000 poorer in 2012 than he or she would have been without the additional regulations. 37 Finally, Wallison invokes public choice theory, which maintains that regulators have their own interests that they seek to advance, sometimes at the expense of the public good for which they are theoretically acting. 38

Wallison acknowledges that some commentators think there is “little alternative” to our strong administrative state because of the complexity of the American economy. 39 But, he replies, that avoids the fundamental question of whether agencies have unconstitutionally arrogated powers to themselves and thereby imperiled our liberties. 40 Wallison also acknowledges the concern of some writers that agencies must enact regulations to address problems in our dynamic economy because Congress is a “cumbersome body that moves slowly in the best of times.” 41 Wallison replies that this response fails to consider how the public wants to be governed and whether agencies have either the legal or moral authority to prescribe rules for society. 42

II. HOW CONGRESS’ DOMINANCE OF OUR NATIONAL GOVERNMENT HAS RECeded

Congress dominated the federal government from the nation’s founding through the Civil War period, although Wallison acknowledges that President Lincoln assumed broad powers to conduct that war. 43 Wallison also argues that some decline in congressional influence occurred even earlier in the nineteenth century, when Martin Van Buren organized slates of Democratic Party members as electors pledged to vote for him for president, thereby creating a new organizational principle that undermined the role of the House of the Representatives in the electoral system. 44 But Congress’ independence was undermined more severely in the twentieth century when presidents like Franklin D. Roosevelt and Lyndon B. Johnson benefitted from landslide election victories and then enlisted the help of Democratic-controlled Congresses to enact their economic and social welfare programs. 45 Party loyalties in Congress, Wallison observes, have “easily overcome” the Framers’ understanding that Congress would be independent of the president and vice versa. 46

The erosion of congressional independence means that Congress does not reliably hold the presidency—including executive branch agencies—accountable, and that during periods of one-party power, the agencies can expand their reach and discretion unhindered. 47 Accordingly, Wallison believes that it is unrealistic for conservatives to expect Congress to restrict agencies’ exercise of authority. 48 Wallison acknowledges that the Trump Administration has used the Congressional Review Act to overturn some regulations promulgated during the Obama Administration, but he is skeptical that the Act is a sufficiently robust vehicle to

29 Id. at 65-68.
30 Id. at 68-70.
31 Id. at 72-74.
32 Id. at 77-108.
33 Id. at 81.
34 Id. at 79.
35 Id. at 86-88, 90-91.
36 Id. at 102-105.
38 Id. at 101.
restrict agency power. Wallison also acknowledges that various bills have been introduced that would require “major” rules to be referred to Congress for approval or disapproval, but he notes that, to date, those bills have not yet been enacted into law.

III. The Demise of the Nondelegation Doctrine and Hope for Its Revival

At the center of Judicial Fortitude is an examination of the nondelegation doctrine and the serious problems posed to our constitutional government by its decline over time. The doctrine “mandate[s] that Congress generally cannot delegate its legislative power to another Branch.” That is because Article I of the Constitution vests the authority to enact legislation exclusively in Congress. Wallison traces this doctrine back to John Locke’s Second Treatise of Government, which influenced the nation’s founders to ground the Constitution in the people’s transfer or delegation to the government of their inherent and natural right to govern themselves. The people’s delegation of lawmakers’ authority could not be subdelegated by Congress absent a constitutional amendment because Congress is the agent of the people and cannot exceed its principal’s instructions. That restriction is consistent with James Madison’s concern, expressed in Federalist No. 47, that separation of powers is necessary to avoid the “tyranny” caused by the consolidation of powers in one of the branches.

Because of the “exclusive nature” of the Constitution’s grant to Congress of “all legislative power,” some legal scholars believe that it is a violation of the Constitution when Congress transfers or delegates any of its legislative authority to administrative agencies. Other scholars conclude that some delegation is inevitable given the complexities of our modern society. But Wallison responds that there is “very little evidence today” that it is necessary to accept broad delegations of congressional authority. Whatever necessity might have existed during the New Deal or Progressive eras to address the country’s “unprecedented” problems, our experiences with many failures of agency governance require at least “substantial evidence” that our constitutional structure cannot manage our contemporary problems just as well, and without the danger of abuse that comes with unaccountable and consolidated government power.

Wallison contends that the nondelegation doctrine “protects and preserves” Congress’ responsibility to make the “most important” decisions for society. If Congress were permitted to delegate that exclusive authority to administrative agencies, the separation of powers would be a “nullity,” and the dangers to our liberty that the Framers feared would become a reality. Major decisions affecting society would be made by the “unelected bureaucracies of the administrative state,” not the people’s representatives.

The courts, however, have not rigorously applied the nondelegation doctrine in recent decades; in fact, the Supreme Court has not applied the doctrine to invalidate an agency action since 1935. In Panama Refining Co. v. Ryan, decided that year, the Court held that a provision of the National Industrial Recovery Act (NIRA) that granted the president power to prohibit the sale of certain oil products constituted an improper delegation of legislative power. In the same year’s A.L.A. Schechter Poultry Corporation v. United States, the Court struck down a different NIRA provision that authorized the president to establish “codes of fair competition,” reasoning that Congress’ failure to define “fair competition” rendered the provision an improper delegation of legislative power. Nine years later, in Yakus v. United States, the Court—by a six to three vote—upheld the authority of the Price Administrator of the wartime Office of Price Administration to set maximum prices for commodities and rents throughout the country at a “generally fair and equitable” level to effectuate the statute’s objectives. The Court found that “Congress has stated the legislative objective, has prescribed the method of achieving that objective,” and has “laid down standards to guide the administrative determination in exercising the delegated authority.” It distinguished Schechter Poultry as a case in which no standards had been defined and the development of prices had been delegated to private entities.

In 2001, in Whitman v. American Trucking Association, the Court articulated its understanding of the nondelegation doctrine’s continued relevance: when Congress confers decision making authority upon agencies, Congress must “lay down by legislative act an intelligible principle to which the person or body...
authorized to [act] is directed to conform.’” Wallison describes this intelligible principle test as “meaningless,” and he contrasts it with Chief Justice Marshall’s articulation of the boundaries of Congress’ Article I authority in Wayman v. Southard, where he drew a line separating “those important subjects which must be entirely regulated by the legislature itself” from those of less interest in which a general provision may be made and power given to those who are to act under such general provisions to fill up the details.” Wallison restates the holding of those who are to act under such general provisions to fill up the in which a general provision may be made and power given to make the “important” decisions, but it can delegate the details of their execution to the agencies or the courts, with the exercise of discretion “confined by the general terms set by Congress.” The modern nondelegation doctrine has strayed far from this principle.

Wallison opines that, although the nondelegation doctrine is not dead, little of its substance remains in modern Supreme Court jurisprudence. He criticizes the Supreme Court for failing to challenge legislation that “hands open-ended legislative authority” to agencies, calling this a “serious failure” of the judiciary to perform its constitutional duties. He contends that the failure to apply a vigorous nondelegation doctrine is the single most important reason for the administrative state’s uncontrolled growth, for the judiciary thereby enables Congress to delegate to agencies broad rulemaking powers.

Wallison also summarizes the views of various legal commentators who have explored the rationale and validity of the nondelegation doctrine, including writings by supporters of the administrative state. He disagrees with scholars who contend that the president can function as an independent check on administrative agencies. Wallison states that it is “fanciful” to think the president alone has the capacity to monitor the thousands of rules issued by agencies each year, and that even then a president cannot “legitimize” an unconstitutional delegation of authority.

What then is the future of the nondelegation doctrine? Wallison detects some “stirrings” at the Supreme Court that indicate that some Justices would like to reexamine the intelligible principle test and attempt a reformulation of the doctrine. Justice Clarence Thomas, for example, remarked in his Whitman concurrence that, in a future case, he would be willing to consider “whether our delegation jurisprudence has strayed too far from our Founders’ understanding of separation of powers.” In a subsequent case, Justice Thomas observed that the Supreme Court had “too long abrogated our duty to enforce the separation of powers required by our Constitution” and had sanctioned an expansive administrative state that makes and enforces laws without accountability. Wallison also notes that both Chief Justice John Roberts and Justice Samuel Alito have expressed concerns about the broad reach of the administrative agencies, and that then-Judge Neil Gorsuch expressed concern about applying the nondelegation doctrine in a Tenth Circuit case.

Justice Brett Kavanaugh also has said that, when an agency wants to exercise expansive regulatory authority over some major social or economic activity, Congress “must clearly authorize” the agency to take such a “major regulatory action.” Wallison concludes that, until the Supreme Court undertakes the task of reviewing the nondelegation doctrine, the doctrine remains “in limbo.” He warns that, so long as Congress continues to enact “goals-oriented” legislation, the courts will continue to confront the nondelegation problem, and he urges the courts to make Congress accountable for its constitutional responsibilities.

IV. Should the Chevron Doctrine be Abrogated?

Wallison is also concerned by the way the Chevron doctrine has expanded agency authority: he characterizes Chevron as an

---

70 Wallison at 116-17, 121, 128 (citing Wayman v. Southard, 23 U.S. 1, 42-45 (1825)).
71 Id. at 119-20.
72 Id. at 126-28. Wallison notes that, as early as J.W. Hampton, Jr. & Co. v. United States, 276 U.S. 394, 409 (1928), the Court “took a major turn away” from the Wayman decision by articulating the “intelligible principle” test. Wallison at 120-21.
73 Wallison at 134.
74 Id.
75 Id. at 128-33.
77 Id. at 128-32, 132-33.
78 Id. at 134-36.
79 Whitman, 531 U.S. at 487 (Thomas, J., concurring).
83 Wallison at 136 (citing United States Telecomm. Ass’n v. FCC, 855 F.3d 381, 421 (2017) (Kavanaugh, J., dissenting from denial of petition for rehearing en banc)).
84 Id.
85 Id.
“easy delegation doctrine”\textsuperscript{87} that has resulted in law- and policy-making by unelected administrators.\textsuperscript{88} Wallison does not appear to disagree with step one of the \textit{Chevron} analysis—in which courts determine whether Congress has “directly spoken to the precise question at issue”—but he disagrees with the Court’s subsequent statement that, if a congressional delegation to an agency on a particular question is “implicit,” the reviewing court cannot substitute its own construction of the statutory provision for a reasonable interpretation made by the agency.\textsuperscript{89} Wallison says that this enables agencies to “infer powers that Congress has not explicitly granted as long as that inference is ‘reasonable.’”\textsuperscript{90} Wallison also criticizes portions of \textit{Chevron} in which the Court indicates that agencies may resolve ambiguities in a statute that have resulted from Congress’ failure to resolve an issue due to competing interests, as well as the Court’s observation that agencies, while not “directly accountable to the people,” are accountable via their supervision by the elected president.\textsuperscript{91}

Wallison rejects the “fiction” that agency decisions are inevitably the consequence of presidential or administration policy, and he contends that the Court has improperly downgraded the role of Congress in deciding policy questions by requiring deference to agency authority when Congress has not resolved an issue.\textsuperscript{92} This wide deference is a “virtual nullification of the separation of powers,” under which Congress becomes merely a source of powers for the agencies rather than a legislative body that decides policy issues for our nation.\textsuperscript{93} Wallison argues that it is Congress’ exclusive role to reconcile conflicting interests and determine resulting policies, that such policies should be reviewable by courts only on constitutional grounds, and that it is not the role of the “federal bureaucracies” in Washington to substitute for Congress’ authority as legislator.\textsuperscript{94} \textit{Chevron}, he contends, permits agencies to displace Congress in our constitutional structure.\textsuperscript{95} Finally, insofar as \textit{Chevron} means that a court must defer to the agency’s interpretation, that deference compromises the court’s independent judgment, “biasing” it in favor of the agency position.\textsuperscript{96} Such deference is inconsistent with the court’s obligation, emphasized by Chief Justice Marshall in \textit{Marbury v. Madison}, to state “what the law is.”\textsuperscript{97}

Wallison recalls Chief Justice Marshall’s statement in \textit{Wayman v. Southard} that Congress is to decide “important issues,” and he argues that agencies should only make policy decisions that are “not important enough” to be made by Congress; even then, he argues, agencies should only make such minor decisions when authorized by Congress to do so.\textsuperscript{98} While recognizing that the Supreme Court has yet to reexamine the \textit{Chevron} doctrine, Wallison discerns in some Justices’ statements a willingness do so, just as those Justices have expressed concern about the decline of the nondelegation doctrine.\textsuperscript{99}

V. IS JUDICIAL FORTITUDE THE SOLUTION TO AN EXPANSIVE ADMINISTRATIVE STATE?

Wallison concludes his book by arguing that the judiciary must assume—or resume—its assigned constitutional role of ensuring that our constitutionally mandated separation of powers is maintained.\textsuperscript{100} He contends that the Framers contemplated that the judiciary not only would interpret the Constitution’s language, but would act to preserve our constitutional structure.\textsuperscript{101} Alexander Hamilton declared in \textit{Federalist} No. 78 that the judiciary would be “faithful guardians” of the Constitution and the “citadel of the public justice and the public security,” and he also warned that the judiciary must remain independent and not become unified with the Congress or the president.\textsuperscript{102} Wallison warns that the Framers’ design of a system in which Congress itself legislates will become obsolete unless the courts intervene.\textsuperscript{103} The courts thus need to revitalize the nondelegation doctrine as it existed before the New Deal Supreme Court retreated from that task.\textsuperscript{104} Otherwise, Congress’ role as the “exclusive source” of power may be compromised.\textsuperscript{105}

\textsuperscript{87} Wallison at 137.

\textsuperscript{88} Id. at 140-41. See \textit{Chevron}, 437 U.S. at 843-44 (deferring to the EPA’s construction of a statutory term insofar as the term was ambiguous and the agency’s interpretation was “permissible”).

\textsuperscript{89} Wallison at 138 (citing \textit{Chevron}, 437 U.S at 843-44).

\textsuperscript{90} Id. at 138.

\textsuperscript{91} Id. at 140 (citing \textit{Chevron}, 437 U.S. at 865).

\textsuperscript{92} Id. Wallison notes that only a few “major” rules are brought to the attention of the Office of Management and Budget, an agency that is within the Executive Office of the President, and that over 12,000 regulations are issued during a typical presidential term. Id. at 143.

\textsuperscript{93} Id. at 141.

\textsuperscript{94} Id.

\textsuperscript{95} Id. at 144.

\textsuperscript{96} Id. at 146 (citing Philip Hamburger, Gorsuch’s Collision Course with the Administrative State, N.Y. Times (March 28, 2017)). My review does not examine whether a court’s application of \textit{Chevron} deference materially affects the outcome as resulting in a pro- or anti-agency decision. See Kristin E. Hickman, SOPRA? So What? Chevron Reform Misses the Target Entirely, 14 U. St. Thomas L.J. 580, 590 (2018) (noting the division of opinion on this issue); Christopher J. Walker, Attacking \textit{Axe} and Chevron Deference: A Literature Review, 16 Geo. J. L. & Pub. Pol’y 103, 120-22 (2018) (citing author’s research that identifies the impact of \textit{Chevron} deference as to favorable outcomes for the agency interpretation).

\textsuperscript{97} Wallison at 138-39 (citing \textit{Marbury v. Madison}, 5 U.S. 137, 177 (1803)).

\textsuperscript{98} Id. at 142.

\textsuperscript{99} Id. at 150-58 (citing \textit{City of Arlington}, 569 U.S. at 315 (Roberts, C.J., Kennedy and Alito, JJ., dissenting); Perez v. Mortgage Bankers Ass’n, 575 U.S. ___, 135 S. Ct. 1199, 1221 (2015) (Thomas, J., concurring) (asserting that principles of deference to agency interpretations are inconsistent with the independent judicial decision making embodied in Article III)). See also Gutierrez-Briuela v. Lynch, 834 F.3d 1142, 1154-55 (10th Cir. 2016) (Gorsuch, J., concurring) (expressing concerns about \textit{Chevron} deference).

\textsuperscript{100} Wallison at 147-50, 161.

\textsuperscript{101} Id. at 147.

\textsuperscript{102} Id. at 137, 146, 165 (citing \textit{The Federalist} No. 78 (Alexander Hamilton)).

\textsuperscript{103} Id. at 147.

\textsuperscript{104} Id. at 147, 149.
of legislation for our national government will be undermined, and our liberty will be threatened.\textsuperscript{105}

Wallison also recommends that the Supreme Court revisit the \textit{Chevron} doctrine.\textsuperscript{106} Borrowing again from Chief Justice Marshall’s opinion in \textit{Wayman v. Southard}, Wallison suggests that courts should evaluate an agency’s interpretation of a statute by whether the agency is addressing a “detail” or an “important decision.” Wallison contends, also must engage in a more searching manner when they interpret agency decisions. First, if there is a lower court decision based on insufficient evidence of Congress’ intent, the appellate court can remand the case to obtain a “further determination” of what Congress intended.\textsuperscript{108} Second, if the court confronts a truly ambiguous statute and the agency has tried to reconcile competing interests in its interpretation, the court can state that it wants a “clearer statement” from Congress before the regulation is enforced, or it can remand the regulation to the agency, “requiring it to put the unresolved question before Congress for a vote”; Congress could enact a procedure to enable such requests.\textsuperscript{109} Alternatively, the agency could declare to Congress its intention to adopt a certain interpretation of an ambiguous statute and, if Congress does not act within a specified time period, the agency’s interpretation would be “deemed correct” by the courts.\textsuperscript{110}

Whatever system is adopted, the result would be a “disciplinary system” for Congress, which would avoid issues of improper delegation or inappropriate levels of deference.\textsuperscript{111} If courts were to send disputed issues of statutory interpretation back to Congress, then Congress would decide that it should resolve the ambiguities and avoid both the problems of \textit{Chevron} deference and additional litigation on that question.\textsuperscript{112} Wallison acknowledges that Congress may find it difficult to reach agreement on controversial issues, but he asserts that it is better to endure “gridlock” than to delegate important decisions to unelected agency officers.\textsuperscript{113}

Wallison acknowledges that some conservatives, including the late Justice Antonin Scalia, were advocates of judicial restraint, having witnessed the Supreme Court’s activism in areas of social policy.\textsuperscript{114} He responds that the decisions that concerned Justice Scalia were based on the Court’s interpretation of the Constitution’s “words,” not the Court’s role in preserving the Constitution’s structure.\textsuperscript{115} In addition, the Court engages in limited, “rational basis,” review of Congress’ regulatory statutes, so a “high hurdle” necessarily exists for competing interests to challenge such laws successfully.\textsuperscript{116}

VI. Are Wallison’s Proposed Reforms Realistic or Appropriate?

Wallison’s recommendation that that the federal judiciary “rein in” the administrative state raises several questions. His primary objective is to shift the dynamic of decision making from administrative agencies to Congress through both increased judicial intervention against delegations of congressional authority and revocation of \textit{Chevron}.\textsuperscript{117} But will increased judicial review of agency decisions—to limit agency law- and policy-making—ultimately result in Congress performing that role more effectively by enacting more detailed, less ambiguous legislation? And will increased judicial review of agency decisions invite more legislating from the bench?

In considering the first question, we might also ask whether the courts can reliably apply Chief Justice Marshall’s distinction between “important questions” and “details” when deciding what constitutes an improper delegation of congressional authority. One judge’s “detail” could be another judge’s “important question.” Litigants, through evidence in the rulemaking record, might be able to identify some demarcation, perhaps relying on estimates of economic or social costs or environmental impact. Whatever test is employed, a court must be satisfied that the statute in question actually empowers the agency to act, and it may be reasonable for courts to adopt limiting constructions of statutes if they doubt the existence of that authority.\textsuperscript{118} Congress then would have to step in and decide the issue directly in subsequent legislation. We may get some enlightenment on these questions if the Court addresses the nondelegation problem at issue in the pending \textit{Gundy} case.\textsuperscript{119}

If the court decides that there has been an unlawful delegation, how should the defect be remedied? Wallison recommends a simple remand by the court to the agency, which would limit the extent of judicial intervention and appropriately refer the issue to the agency.\textsuperscript{120} But while that inquiry is pending, will the litigation be held in abeyance? And although Wallison suggests that the agency can ask Congress to vote on the

\textsuperscript{105} Id. at 149.
\textsuperscript{107} Wallison at 158-59.
\textsuperscript{108} Id. at 159.
\textsuperscript{109} Id.
\textsuperscript{110} Id. at 160.
\textsuperscript{111} Id. at 160.
\textsuperscript{112} Id. Wallison expresses some confidence that Congress can address and resolve complex issues, delegating only “technical matters” to the agencies. Id. at 163.
\textsuperscript{113} Id. at 164.
\textsuperscript{114} Id. at 148-49.
\textsuperscript{115} Id. at 149.
\textsuperscript{116} Id. at xxii–xxiii.
\textsuperscript{117} Id. at 111, 134-35.
\textsuperscript{118} See C. Boyden Gray, \textit{The Nondelegation Canon’s Neglected History and Underestimated Legacy}, 22 GEO. Mason L. REV. 619, 623-25, 646 (2015) (recommending that courts monitor potential agency overreach through careful review of purported agency authority and construe statutes narrowly to avoid such overreach).
\textsuperscript{119} See supra note 82.
\textsuperscript{120} Wallison at 159.
unresolved question, will that new process create a “backlog” of pending inquiries? Will Congress prioritize the requests?

With respect to the delegation problem, can we reasonably expect Congress to legislate in detail on matters as diverse as ratemaking, the licensing of broadcast stations, defining unfair or deceptive industry practices, or complex scientific or technical issues involved in safety or environmental regulations? Might a more prudent course be to focus Congress’ oversight function on its reviewing proposed major rules promulgated by agencies before the rules can become effective? These may be more effective—because more feasible—means of policing agency overreach.

Similar problems may arise if the courts abandon Chevron and undertake de novo statutory review of ambiguous statutory provisions. If a court concludes that a statute is ambiguous, will the court resolve the ambiguity, or will it remand to the agency or Congress for additional consideration of the issue? The more appropriate course would be for the court to avoid providing its own interpretation—and thus avoid its own overreach—but to instead demand a better explanation from the agency. But in cases in which Congress has not addressed an inadvertent ambiguity in a statutory provision, the agency on remand may not necessarily have additional insight to provide to the court. In those situations, Congress alone could resolve the ambiguity. But there may be situations in which Congress did not intend to legislate with specificity, purposely relying on the agency’s expertise to address and resolve the question.

The central challenge of separation of powers is to create an appropriate equilibrium between Congress and the executive branch, while ensuring that the judiciary remains the branch “least dangerous to the political rights of the Constitution.” Wallison wants the courts to enforce the Constitution’s structural protections, but he is understandably leery of the courts extending their reach to resolve broader issues. I also am concerned that increased judicial scrutiny of agency interpretations of statutes could result in judges replacing agency administrators as de facto legislators. Judicial ideologies might supplant agency decision making, and while career officials are not at-will employees, they also do not have life tenure like federal judges do. There is some basis to conclude that judges might be freer, absent Chevron, to overreach. Whether our judiciary is progressive or conservative ideologically, are we comfortable with the judiciary determining the legal contours of our civic and economic arrangements? Notwithstanding these concerns, Wallison has framed the nondelegation and Chevron deference issues in an eloquent and reasoned way, and he has made a valuable contribution to the ongoing debate on those doctrines.

121 See Robert R. Gasaway, Ashley C. Parrish, Administrative Law in Flux: An Opportunity for Constitutional Reassessment, 24 Geo. Mason L. Rev. 361, 395-96 (2017) (recommending that Congress “bind itself to taking legislative action (even if the action is an affirmative decision not to act)” in various administrative law contexts, including “fast track” proposals on agency-submitted proposals to modify rules governing exercises of agency discretion, or agency responses to Supreme Court decisions holding legislation unconstitutional on non-delegation grounds).

122 See Nicholas R. Bednar, Kristin E. Hickman, Chevron’s Inevitability, 85 Geo. Wash. L. Rev. 1392, 1454 (2017) (“Congress routinely delegates authority to agencies because it wants to utilize their policy and scientific expertise to resolve programmatic details and fill statutory gaps”).

123 For an example of a pending bill on that topic, see H.R. 26, the Regulations from the Executive In Need of Scrutiny Act, 115th Cong., 1st Sess., https://www.congress.gov/115/bills/hr26/BILLS-115hr26rfs.pdf. The bill states that “major rules” would not have legal effect until Congress passes (and the President signs) a joint resolution approving the rules.

124 For an example of a pending bill on that topic, see H.R. 76, the Separation of Powers Restoration Act of 2017, 115th Cong., 1st Sess., https://www.congress.gov/bill/115th-congress/house-bill/76/text. The bill would amend section 706 of the Administrative Procedure Act, 5 U.S.C. § 706, to provide that courts would review all relevant statutory questions de novo. See Bednar and Hickman, supra note 122, at 1456-58 (critiquing these proposed reforms).

125 See Richard J. Pierce, The Future of Deference, 84 Geo. Wash. L. Rev. 1293, 1313 (2016) (the elimination of deference doctrines would come at a “very high cost” in terms of the failure to recognize the agencies’ “superior knowledge of the field and superior understanding of the ways in which an interpretation of a statute affects the ability of the agency to implement a coherent and efficient regulatory regime”).

126 See The Federalist No. 78 (Alexander Hamilton).

127 Wallison at 82, 147-19.

128 See Kent Barnett, Christina L. Boyd, and Christopher J. Walker, Administrative Law’s Political Dynamics, 71 Vand. L. Rev. 1463, 1467, 1524 (2018) (describing empirical research on the application of Chevron to over 1,500 circuit court cases over 11 years (2003-2013) and concluding that Chevron deference “has a powerful constraining effect on partisanship in judicial decisionmaking”).
On September 27, 2018, the Supreme Court granted certiorari in Rimini Street, Inc. v. Oracle USA, Inc., to address “whether the Copyright Act's allowance of 'full costs' (17 U.S.C. § 505) to a prevailing party is limited to taxable costs under 28 U.S.C. §§ 1920 and 1821 . . . or also authorizes non-taxable costs . . . .” The case raises a textbook statutory interpretation issue, and it illuminates common pitfalls in construing statutes that are particularly tempting for textualists.

I. BACKGROUND

28 U.S.C. § 1920 (Section 1920) is “the general statute governing the taxation of costs in federal court.” The law allows federal courts to “tax as costs” a specific set of expenditures listed in the statute. These are: “(1) Fees of the clerk and marshal; (2) Fees for printed or electronically recorded transcripts necessarily obtained for use in the case; (3) Fees and disbursements for printing and witnesses; (4) Fees for exemplification and the costs of making copies of any materials where the copies are necessarily obtained for use in the case; (5) Docket fees under section 1923 of this title; [and] (6) Compensation of court appointed experts, compensation of interpreters, and salaries, fees, expenses, and costs of special interpretation services under section 1828 of this title.” 28 U.S.C. § 1920.

4 These are: “(1) Fees of the clerk and marshal; (2) Fees for printed or electronically recorded transcripts necessarily obtained for use in the case; (3) Fees and disbursements for printing and witnesses; (4) Fees for exemplification and the costs of making copies of any materials where the copies are necessarily obtained for use in the case; (5) Docket fees under section 1923 of this title; [and] (6) Compensation of court appointed experts, compensation of interpreters, and salaries, fees, expenses, and costs of special interpretation services under section 1828 of this title.” 28 U.S.C. § 1920.

will be construed as authorizing” additional costs beyond those enumerated therein “unless the statute refers explicitly” to them.6

17 U.S.C. § 505 (Section 505), part of the Copyright Act of 1976 (but which dates back to 1831), provides that:

In any civil action under this title, the court in its discretion may allow the recovery of full costs by or against any party other than the United States or an officer thereof. . . . the court may also award a reasonable attorney’s fee to the prevailing party as part of the costs.7

Notably for this case, this provision allows for the award of “full costs,” not just “costs.”

The question in this case is straightforward: does “full costs” in Section 505 expand the scope of awardable expenditures beyond what Section 1920 permits?8 The trial court in this case said “yes,” concluding that “full costs” means “all costs incurred in litigation,” not just those identified in Section 1920.9 Accordingly, the court awarded Oracle millions of dollars in “non-taxable costs,” including “litigation costs for expert witness fees.”10 The Ninth Circuit affirmed in relevant part.11 Following controlling circuit precedent, the panel reasoned that reading “full costs” as only those costs enumerated in Section 1920 “effectively reads the word ‘full’ out of the statute,” and that such a construction violates the canon against surplusage that requires a court to “give effect, amount . . . for something” or “the outlay or expenditure . . . made to achieve an object.”12 Given courts’ duty “to give effect, if possible, to every . . . word of a statute,”13 the Ninth Circuit’s reasoning is superficially attractive: “full costs” permits recovery of “all costs incurred in litigation,” not just Section 1920 costs. From Oracle’s point of view, this is simply taking Section 505 to “mean[] what it says.”14

As the Court’s grant of certiorari suggests, however, things are not so simple. “[S]tatutory interpretation,” of course, “is a ‘holistic endeavor,’” and “the words of a statute are not to be read in isolation.”15 Rather, terms in a statute must be “read together” with the rest of the law.16 The “plain meaning” of a given statutory provision must be ascertained not just from “the language itself,” but also from “the specific context in which that language is used, and the broader context of the statute as a whole.”17 This is a “fundamental principle,” not only of “statutory construction” but also “of language itself.”18 Thus, even reading for “plain meaning,” one cannot simply stop at the “dictionary definition of two isolated words” in trying to figure out what the law means; rather, statutory interpretation must be conducted in light of “the text and structure” of the law as a whole.19

Additionally, “it is an established rule of law, that all acts in pari materia are to be taken together, as if they were one law.”20 Under this canon of statutory construction, statutes addressing the same subject matter generally should be read as if consisting of one law addressing the subject.21 Laws so related to one another thus constitute part of the “broader context” to be taken into consideration when ascertaining the plain meaning of statutory terms.22

II. Reading Law: A “Holistic Endeavor”

At first blush, this case might seem simple, particularly to a textualist. Per the dictionary, “full” means “containing as much . . . as is possible” or “complete;,”23 and “cost” means “the amount . . . for something” or “the outlay or expenditure . . . made to achieve an object.”24 Given courts’ duty “to give effect, if possible, to every . . . word of a statute,” the Ninth Circuit’s

---

6  Murphy, 548 U.S. at 301.
8  See Humphreys & Partners Architects v. Lessard Design, 152 F. Supp. 3d 503, 524 (E.D. Va. 2015) (“Section 505 allows a court to award ‘full costs’; it is unclear, however, what ‘full costs’ means.”).
10  Id.
11  Oracle USA, Inc. v. Rimini Street, Inc., 879 F.3d 948, 965–66 (9th Cir. 2018).
12  Twentieth Century Fox Film Corp. v. Entm’t Distrib., 429 F.3d 869, 885 (9th Cir. 2005).
22  United States v. Freeman, 44 U.S. (3 How.) 556, 564 (1845).
24  See Erlenbaugh v. United States, 409 U.S. 239, 243–44 (1972) (cleaned up) (“The rule of in pari materia—like any canon of statutory construction—is a reflection of practical experience in the interpretation of statutes: a legislative body generally uses a particular word with a consistent meaning in a given context. Thus, for example, a later act can . . . be regarded as a legislative interpretation of an earlier act . . . in the sense that it aids in ascertaining the meaning of the words as used in their contemporary setting, and is therefore entitled to great weight in resolving any ambiguities and doubts. The rule is but a logical extension of the principle that individual sections of a single statute should be construed together, for it necessarily assumes that whenever Congress passes a new statute, it acts aware of all previous statutes on the same subject.”).
Another relevant principle is the “commonsense of statutory construction that the specific governs the general.”25 Accordingly, the Court will not read a general clause in one place in a way that undermines a carefully drawn statute elsewhere.26 This principle holds no matter how “inclusive may be the general language of a statute,” particularly when “Congress has enacted a comprehensive scheme and has deliberately targeted specific problems with specific solutions.”27

In addition to the foregoing, it must be remembered that some statutes use terms of art, in which case the plain meaning—understood as the dictionary definition—is less relevant.28 Sometimes, for instance, Congress may use language that comes freighted with meaning due to prior constructions of that language.29 Other times, Congress may employ true terms of art that must be understood according to their technical, not plain, meaning.30 In either case, when Congress writes laws with such words is not apt to illuminate its meaning.31

III. The Law Means What It Says, But What Does It Say?

While “full costs” of course “means what it says,” as Oracle contends, dictionary definitions alone do not necessarily tell us what it says.32 Construing those words in the context of federal “costs” laws generally, and in light of the grammatical-historical background to the phrase “full costs,” it turns out that, as Rimini Street argues, “full costs” means all costs enumerated in Section 1920.33

First, consider the larger statutory context. The Supreme Court has found that Section 1920 “embodies Congress’ considered choice as to the kinds of expenses that a federal court

28 FAA v. Cooper, 566 U.S. 284, 291–92 (2012) (noting that when Congress uses terms of art, “the ordinary meaning of the word[s]” as “defined in standard general-purpose dictionaries” is not dispositive, but rather the interpreter must take into account “the cluster of ideas” incorporated into the special term); see also Molzof v. United States, 502 U.S. 301, 306–07 (1992); cf. Yates v. United States, 354 U.S. 298, 319 (1957), overruled on other grounds by Burks v. United States, 437 U.S. 1 (1978) (“[W]e should not assume that Congress . . . used . . . words . . . in their ordinary dictionary meaning when they had already been construed as terms of art carrying a special and limited connotation.”).
29 See, e.g., Hall v. Hall, 138 S. Ct. 1118, 1125 (2018) (“This is not a plain meaning case. It is instead about a term . . . with a legal lineage stretching back at least to the first federal consolidation statute, enacted by Congress in 1813. Over 125 years, this Court . . . interpreted the term ‘full costs’ to include only those expenditures listed in 28 U.S.C. §§ 1821 and 1920.”).
30 Cooper, 566 U.S. at 291–92.
32 See Oracle Brief at 1 (“When the text of a statute is clear, judicial inquiry ends where it begins—with the text.”).
33 See Farmcy Records v. Nassar, 729 F. Supp. 2d 865, 893 (E.D. Mich. 2010) (“The majority of courts which have considered the issue have interpreted the term ‘full costs’ to include only those expenditures listed in 28 U.S.C. §§ 1821 and 1920.”).
34 Crawford Fitting, 482 U.S. at 440.
35 Id. at 441. See Humphreys & Partners Architects, 152 F. Supp. 3d at 524 (recognizing that Crawford Fitting “provided a general framework for considering whether a statute provides for the recovery of costs that exceed the scope of costs recoverable under Sections 1821 and 1920”).
36 Crawford Fitting, 482 U.S. at 442.
38 284 U.S. 444, 444–45 (1932).
39 Id. at 446.
41 Crawford Fitting, 482 U.S. at 441.
loaded with meaning because of its grammatical-historical background in U.S. law. As Rimini Street points out, when the phrase “full costs” first appeared in U.S. copyright law, Congress was legislating against a background where a prevailing party might not be awarded any costs whatsoever.43 In response to this, Congress passed a law mandating that “full costs shall be allowed.”44 At the time, federal courts followed the forum state’s law with respect to costs, and so “full costs” in the copyright law effectively meant the total amount of whatever costs were allowed under state law, which could vary from state to state.45 In 1853, however, Congress enacted the Fee Act, “specifying in detail the nature and amount of the taxable items of cost in the federal court.46 At the time, federal courts followed the forum state’s law, which could vary from state to state.45 In 1853, however, Congress enacted the Fee Act, “specifying in detail the nature and amount of the taxable items of cost in the federal courts”46 and providing that “no other compensation shall be allowed.”47 The Fee Act thus replaced the previously-controlling patchwork of state laws and became the relevant law defining the “full costs” awardable in federal court, including in cases involving federal copyright law. The language of “full costs” was retained over time in the copyright law, while the Fee Act laid the foundation for current Section 1920.48 Accordingly, “full costs” now means those costs enumerated in Section 1920, which details the total amount of costs allowable as a matter of course in federal court.49

IV. But Mustn’t “Full” Mean “Full”?

Oracle, however, disagrees. In its briefing, Oracle focuses particularly on plain meaning and the canon against surplusage, but its arguments prove unavailing.

A. “Full” Is Not “Plain Evidence of Congressional Intent”

Oracle argues that the word “full” loosens the scope of “costs” to allow for recovery of expenses beyond those delineated in Section 1920.50 This argument runs up against Congress’ clear authority to allow for the award of costs beyond those taxable costs normally awarded a prevailing party under § 1920. The Court disagrees with Devcom’s interpretation. Caselaw interpreting 17 U.S.C. § 116 (1970) (repealed), the precursor of § 505, has never accorded courts more discretion because of the word ‘full.’

51 See, e.g., Casey, 499 U.S. at 89 (noting how “[a]t least 34 statutes in 10 different titles of the U.S. Code explicitly shift attorney’s fees and expert witness fees”).

52 See Epic Sys. Corp. v. Lewis, 138 S. Ct. 1612, 1626 (2018); Crawford Fitting, 482 U.S. at 445; see Casey, 499 U.S. at 86 (noting that Sections 1920 and 1821 “define the full extent of a federal court’s power to shift litigation costs absent express statutory authority to go further”); cf. Fogerty v. Fantasy, Inc., 510 U.S. 517, 524 n.11 (1994) (“The 1976 Copyright Act did change . . . the standard for awarding costs to the prevailing party. . . . The 1976 Act changed the rule from a mandatory one to one of discretion. As the 1909 Act indicates, Congress clearly knows how to use mandatory language when it so desires.”).

53 Crawford Fitting, 482 U.S. at 445. See Casey, 499 U.S. at 86 (noting that these provisions “define the full extent of a federal court’s power to shift litigation costs absent express statutory authority to go further”).

54 Humphrey’s & Partners Architects, 152 F. Supp. 3d at 524–25.

55 Crawford Fitting, 482 U.S. at 445. See BMG Rights Mgmt. (US) v. Cox Commun’cs, 234 F. Supp. 3d 760, 779 (E.D. Va. 2017) (“Although there is reasonable debate over the proper interpretation of the word ‘full’ in § 505, that term is certainly not explicit in authorizing witness fees or any other non-taxable costs. If it wanted to, Congress could have easily inserted language allowing for the recovery of any category of costs, however, it chose not to, and the Court will not implicitly read those terms into the statute.”); Tempest Pub’g, Inc. v. Hacienda Records & Recording Studio, Inc., 141 F. Supp. 3d 712, 723 (S.D. Tex. 2015) (“Although there is some evidence that Congress intended costs recoverable under § 505 to exceed those recoverable under § 1920, that evidence is not clear or explicit, as needed to conclude that the general statute controls the more specific one. Crawford Fitting, 482 U.S. at 445. When the case is close, Crawford Fitting’s standard counsels in favor of following the specific statute and applying the presumption against implied repeals. Id. Given that standard, the Supreme Court’s opinion in Marx, and the weight of circuit authority resting against a broad reading of § 505, the court concludes that the costs taxable under § 505 are limited to those enumerated in § 1920.”).

about whether attorneys’ fees could be recovered as costs . . . .”\(^{57}\)

According to Oracle, “Congress added the second sentence to the Copyright Act to clarify the continued availability of attorneys’ fees” as part of the “full costs” available under Section 505.\(^{58}\) In response to the retort that this provision would be superfluous if “full costs” already included attorneys’ fees, Oracle explains that “clarity is not ‘superfluous.’”\(^{59}\) That may be true, but it raises another problem for Oracle: Congress singled out attorneys’ fees to make them taxable as costs, but it did not do the same for expert witness fees, so the taxability of attorneys’ fees actually works against Oracle’s position.\(^{60}\) The 1909 amendment shows that Congress knows how to make its intention clear as to recoverable costs in Section 505 when it wants to permit costs beyond those allowed via Section 1920.\(^{61}\) Given that the circuit split at issue here is not new,\(^{62}\) Congress “could easily have” amended the statute again to indicate that expert witness fees are also recoverable as part of the costs.\(^{63}\) Congress has not done so, instead leaving Section 505 to allow only “full costs” and “a reasonable attorney’s fee.” The explicit grant of authority to award one type of fee “as part of the costs” implies that a court may not award any other type of fee beyond normally allowable costs.\(^{64}\)

\(^{57}\) Oracle Brief at 5.

\(^{58}\) Id. at 14.

\(^{59}\) Id. at 29.

\(^{60}\) See \textit{Casey}, 499 U.S. at 95 (“Congress . . . having authorized the taxation of reasonable attorney’s fees without making any provision with respect to . . . fees of expert witnesses, must presumably have intended that they not be taxed.”) (internal quotation marks omitted).

\(^{61}\) See \textit{Fordyce}, 510 U.S. at 524 n.11 (“As the 1909 Act indicates, Congress clearly knows how to use mandatory language when it so desires.”); see also \textit{Casey}, 499 U.S. at 100–01 (responding to the argument that Congress would have included expert fees in a fee-shifting statute “had it thought about it” by explaining that “[t]he facile attribution of congressional ‘forgetfulness’ cannot justify” departing from “that permissible meaning which fits most logically and comfortably into the body of both previously and subsequently enacted law”); cf. \textit{Transamerica Mortg. Advisors (tama) v. Lewis}, 444 U.S. 11, 22 & n.13 (1979) (noting that “[w]hile subsequent legislation can disclose little or nothing of the intent of Congress in enacting earlier laws,” statutory amendments at least reveal “that Congress knew how to” overcome interpretive presumptions—in this case, the presumption against implied private rights of action—“when it wished to do so,” so Congress’ failure to overcome such presumptions is evidence that the presumption should still control).

\(^{62}\) As Oracle argued in opposition to certiorari, the circuit split on this issue is “stale”—the circuit court decisions holding that “full costs” only means “taxable costs” are over fifteen years old. Oracle Opposition at 2, 14.

\(^{63}\) \textit{Casey}, 499 U.S. at 99.

\(^{64}\) The well-established rule of expressio unius supports this conclusion. See, e.g., \textit{Jennings v. Rodriguez}, 138 S. Ct. 830, 844 (2018).

\(^{59}\) Note that “full” in Section 505 should be considered a “delocalized” adjective, which is an adjective “whose purpose is to draw attention to and underline an attribute that is already embedded in the meaning of the noun . . . .”

\(^{65}\) \textit{Marx v. General Revenue Corporation}, for example, the Court rejected a surplusage argument against reading a costs statute in a certain way.\(^{66}\) The \textit{Marx} Court reasoned that “[t]he canon against surplusage is not an absolute rule,” as “‘instances of surplusage are not unknown,’” and “redundancy is ‘hardly unusual’ in statutes addressing costs.”\(^{67}\) In fact, the Court here hinted that the canon is effectively of no use at all in the costs context since “a court has inherent power” to shift costs in some cases, meaning that there is “no need for Congress to specify that courts have this power” in those circumstances, making various federal statutes to this effect wholly superfluous.\(^{68}\)

Second, Oracle’s insistence on plain or ordinary meaning conflicts with its argument against surplusage. The adjective “full” quite frequently is redundant in its ordinary usage, serving only to emphasize or clarify meaning already inherently contained within the modified noun.\(^{69}\) For instance, imagine an automobile driver requesting that a passenger help defray transportation costs by buying fuel. There is no difference between asking the passenger to pay for “a tank of gas” or “a full tank of gas.” Either way, the passenger (if she is polite) is going to fill the tank. “Tank” alone means the same thing as “full tank.” The use of “full” simply emphasizes and makes clear what the word “tank,” standing alone, already indicates. Thus, the plain meaning and ordinary use of

\(^{66}\) See \textit{Oracle Opposition} at 18 (“Petitioners attempt to resist that conclusion by positing that ‘full’ simply means that a prevailing party can recover the entirety of the costs allowable under §1920. But they do not and cannot explain why Congress would need to include that clarification since the same is true under §1920 itself. Indeed, statutes often authorize recovery of ‘costs simpliciter’ without specifying that each enumerated cost may be recovered ‘in full.’”).

\(^{67}\) See \textit{Oracle Opposition} at 18 (“Petitioners attempt to resist that conclusion by positing that ‘full’ simply means that a prevailing party can recover the entirety of the costs allowable under §1920. But they do not and cannot explain why Congress would need to include that clarification since the same is true under §1920 itself. Indeed, statutes often authorize recovery of ‘costs simpliciter’ without specifying that each enumerated cost may be recovered ‘in full.’”).

\(^{68}\) \textit{Id.}

\(^{69}\) \textit{Id.}

\(^{69}\) See Brief of Amici Curiae Scholars of Corpus Linguistics Supporting Petitioners, \textit{Rimini Street, Inc. v. Oracle USA Inc.}, No. 17-1625, at 2 (Nov. 20, 2018), \url{https://www.supremecourt.gov/DocketPDF/17/1625/72865/20181120195011397_Rimini%20Street%20Amicus%20FINAL%20pdf.pdf} (“‘Full’ in Section 505 should be considered a ‘delocalized’ adjective, which is an adjective ‘whose purpose is to draw attention to and underline an attribute that is already embedded in the meaning of the noun . . . .’”).
“full” is such that it is often literally superfluous.\textsuperscript{70} And, when “ordinary meaning would render [a] term superfluous,” the canon against surplusage “should not be used to distort [that] ordinary meaning,” a point that is especially true in “obvious instances of iteration to which lawyers . . . are particularly addicted.”\textsuperscript{71}

In support of its surplusage argument, Oracle points to four federal statutes—all enacted after the 1976 passage of the Copyright Act—that allow for recovery of “full costs,” arguing that Rimini Street’s reading would “render superfluous the word ‘full’ in all four.”\textsuperscript{72} On the contrary, those statutes support the argument that “full costs” does not mean “all costs incurred in litigation.”\textsuperscript{73} Take, for instance, the Digital Millennium Copyright Act of 1998, which states that a “court in its discretion may allow the recovery of full costs by or against any party and may also award a reasonable attorney’s fee to the prevailing party as part of the costs.”\textsuperscript{74} If “full costs” generally means all litigation-related expenses, including costs like attorneys’ fees and expert witness expenses, Congress did not need to say that the court could award “full costs” and “also . . . a reasonable attorney’s fee.” Whether or not Section 505 needed “clarifying” in the 1909 amendments,\textsuperscript{75} there is no reason why, nearly a century later, Congress would still be clarifying that courts may award “reasonable attorney’s fees[s] to the prevailing party as part of the costs” if “full costs” included such expenses. To the same effect is the Semiconductor Chip Protection Act of 1984, which provides that “the court in its discretion may allow the recovery of full costs, including reasonable attorneys’ fees, to the prevailing party.”\textsuperscript{76} The clause in this statute regarding attorneys’ fees is superfluous if “full costs” already includes them. Yet again, the Cable Communications Policy Act of 1984 provides for “the recovery of full costs, including . . . reasonable attorneys’ fees to an aggrieved party who prevails.”\textsuperscript{77} If “full costs” already included attorneys’ fees, the latter half of this provision is superfluous. In short, if the canon against surplusage still applies here given the risk that “multiple sentences in multiple enactments would be rendered nugatory,”\textsuperscript{78} the canon militates in favor of Rimini Street’s position: “full costs” generally includes all litigation-related expenses except attorneys’ fees. But of course, it is also an interpretive presumption that the term “costs” is “defined by the categories of expenses enumerated in 28 U.S.C. § 1920,” and the Court has held that “no statute will be construed as authorizing the taxation of witness fees as costs unless that statute refers explicitly to witness fees.”\textsuperscript{79} In other words, if clear language is needed to overcome the American Rule prohibiting an award of attorneys’ fees as costs, so too clear language is needed to overcome the presumption that witness fees are not part of awardable costs. Section 505 overcomes the American Rule by allowing for an award of attorneys’ fees, but Section 505 is silent with respect to other non-taxable litigation expenses and so does not overcome what could be dubbed the “Crawford Fitting Rule” prohibiting an award of litigation expenses beyond Section 1920 absent a clear statement allowing them.

C. The Limits of Precedent

Oracle cites Ninth Circuit precedent and cases from other circuits to support its position.\textsuperscript{81} But the reasoning of those cases either does not hold up or does not in fact support Oracle. There is no question that Twentieth Century Fox supports Oracle’s position. In that case, the Ninth Circuit relied on the canon against surplusage to read “full costs” as allowing for recovery of expenses beyond those delineated in Section 1920.\textsuperscript{82} However, the other authorities upon which Oracle relies provide little, if any, support for the position that full costs means something more than “the entire amount of allowable costs under Section 1920.”

In Coles v. Wonder, for instance, the Sixth Circuit affirmed an award under Section 505 of “$14,172.34 in non-taxable costs,” but the court focused only on the “attorney’s fee” sentence in the statute in analyzing the validity of the award.\textsuperscript{83} The court affirmed the district court’s entire award only after noting that it owed deference “to the discretion of the district court in the award of attorney’s fees . . . .”\textsuperscript{84} In fact, courts in the Sixth Circuit have explicitly rejected Oracle’s reading of Coles, holding that

\textsuperscript{70} As the Corpus Linguistics amici explain, the “delexicalization” of the word “full” is quite common: a person in possession of a “deck of cards” would be presumed to have a “full deck,” for instance, so “deck” and “full deck” communicate the same point in ordinary parlance. See generally id.

\textsuperscript{71} Moskal v. United States, 498 U.S. 103, 120 (1990) (Scalia, J, dissenting).

\textsuperscript{72} Oracle Brief at 19; see also id. at 23.

\textsuperscript{73} Oracle United States, Inc., 209 F. Supp. 3d at 1218.

\textsuperscript{74} 28 U.S.C. § 4001(g).

\textsuperscript{75} See supra notes 56–64 and accompanying text.

\textsuperscript{76} 17 U.S.C. § 911(f).


\textsuperscript{78} Oracle Brief at 23.

\textsuperscript{79} See, e.g., Oracle Brief at 25.

\textsuperscript{80} Murphy, 548 U.S. at 301 (cleaned up) (emphasis added).

\textsuperscript{81} See Oracle Opposition at 14.

\textsuperscript{82} 429 F.3d 869, 885 (9th Cir. 2005).

\textsuperscript{83} 283 F.3d 798, 803–04 (6th Cir. 2002).

\textsuperscript{84} Id. (emphasis added). See R.C. Ofling, Inc. v. CU Interface, LLC, No. 5:08CV234, 2011 U.S. Dist. LEXIS 15595, at *22–24 (N.D. Ohio Feb. 16, 2011) (awarding non-taxable costs under Section 505 per Coles by reasoning that such costs “may be subsumed within the phrase attorney’s fees” appearing in the statute).
non-taxable costs are not awardable under Section 505.\textsuperscript{85} Furthermore, the Sixth Circuit has repudiated \textit{Coles} to the extent it can be read as allowing costs beyond those enumerated in Section 1920.\textsuperscript{86} The First Circuit in \textit{InvesSys, Inc. v. McGraw-Hill Cos.} found non-taxable costs “recoverable under § 505,”\textsuperscript{87} but it found them recoverable as part of an “attorney’s fee” under the law, not as part of “full costs.”\textsuperscript{88} In fact, the \textit{InvesSys} court recognized “the tendency of the courts” to treat “full costs” in Section 505 just as “costs” elsewhere, and it found that the general consensus is that non-taxable costs are not recoverable as “full costs” because Section 1920 “does not include” them.\textsuperscript{89} Thus, \textit{InvesSys} does not support the view that non-taxable costs may be assessed as part of the “full costs” permitted by Section 505.

Finally, while the Seventh Circuit commented in dicta in \textit{Susan Wakeen Doll Co. v. Ashton-Drake Galleries} that “non-taxable costs” could “come through [Section 505],” the court did not specify whether such costs would “come through” as “full costs” or as “attorney’s fee[s].”\textsuperscript{90} The fact that, in context, the court was addressing the district court’s award of attorneys’ fees suggests that the Seventh Circuit here may have been thinking, like the First Circuit in \textit{InvesSys}, that non-taxable costs could be awarded under Section 505 as part of an award of “attorney’s fee[s].”\textsuperscript{91} In other words, \textit{Susan Wakeen Doll Co.} does not support the idea that “non-taxable costs” may be awarded as “full costs” under Section 505.  

\textbf{D. The Technical Meaning of “Full Costs”}

Finally, according to Oracle, “the original practice in copyright cases was for prevailing parties to receive all the costs they expended in the litigation . . . .”\textsuperscript{92} This was because federal courts followed state law in fashioning awards of costs and fees in copyright suits, and the states reportedly permitted taxation of all litigation expenses, not just what are now referred to as taxable costs.\textsuperscript{93} However, uncertainty crept into the law after an 1819 statute gave federal circuit courts original jurisdiction over copyright cases, calling into question whether state law or federal law should furnish the rule of decision regarding awardable costs in copyright actions. The Copyright Act of 1831, where “full costs” first appeared, thus “reinstated the default state rule” allowing recovery of the full gamut of litigation expenses.\textsuperscript{94} This argument, however, proves Rimini Street’s point: that “full costs” in Section 505 permits recovery of all that is recoverable under the governing law. Whether or not it was the case that the governing law in 1831 allowed for recovery of all litigation-related expenses as “full costs” because of the relevant state laws, Section 1920 is the governing law now, and it plainly does not.

\textbf{V. Barking Up the Wrong Tree}

There is one curious detail about the case, mentioned in passing by Oracle, that reinforces the foregoing. The non-taxable costs Oracle seeks to recover were purportedly incurred due to Rimini Street’s “intentional spoliation of evidence and lying under oath,” which “forced Oracle to expend an extraordinary amount of resources proving conduct” that Rimini Street committed.\textsuperscript{95} Federal courts have “inherent power” to address “[a]llegations of spoliation, including the destruction of evidence in pending or reasonably foreseeable litigation,” and the exercise of this “inherent power” is particularly called for where “there is no statute or rule that adequately addresses the conduct.”\textsuperscript{96} Courts, in fact, have invoked this “inherent authority” in similar situations where “reimbursement of . . . fees and expenses that can be traced to the spoliation” was called for “to remedy the expenses incurred” by a party.\textsuperscript{97} In this case, it appears the trial court invoked this inherent power, sanctioning Rimini Street for spoliation by giving an adverse inference instruction to the jury.\textsuperscript{98} According to Oracle, the trial court also had “ample authority to shift costs as a sanction for . . . misconduct,” but it did not, choosing instead to make an award of non-taxable costs under Section 505 that included some of the expenditures Oracle alleges

\textsuperscript{85} See, e.g., Liang v. AWG Remarketing, Inc., No. 2:14-cv-00099, 2016 U.S. Dist. LEXIS 13566, at *52–54 (S.D. Ohio Feb. 4, 2016) (noting that \textit{Coles} “contains no discussion on this issue” before concluding that there was no reason to conclude that “the Sixth Circuit would adopt the rule adopted by the Ninth Circuit that non-taxable costs . . . are recoverable as costs under the Copyright Act”); \textit{Pharmacy Records}, 729 F. Supp. 2d at 893 (noting that \textit{Coles} “affirmed an award of non-taxable costs under § 505, without discussion,” before reviewing \textit{Crawford Fitting} and holding that “the fees paid by the Defendants to their consulting experts are not recoverable against the Plaintiffs under the Copyright Act”).

\textsuperscript{86} See L & W Supply Corp. v. Acuity, 475 F.3d 737, 738–41 (6th Cir. 2007) (“Witness fees are not recoverable as costs absent explicit statutory authority. . . . any earlier Sixth Circuit and/or any other earlier precedent is no longer controlling.”).

\textsuperscript{87} Oracle Opposition at 14.

\textsuperscript{88} 360 F.3d 16, 22–23 (1st Cir. 2004). \textit{See also} \textit{Tempest Publ’g}, 141 F. Supp. 3d at 722 (recognizing that the \textit{InvesSys} court “held that electronic-research costs are recoverable as part of attorney’s fees”).

\textsuperscript{89} 369 F.3d at 22.

\textsuperscript{90} 272 F.3d 441, 458 (7th Cir. 2001).

\textsuperscript{91} Id. at 457–58.

\textsuperscript{92} Oracle Opposition at 20.

\textsuperscript{93} Id.

\textsuperscript{94} Id. at 21–22.

\textsuperscript{95} Id. at 12.

\textsuperscript{96} Rinkus Consulting Grp., Inc. v. Cammarata, 688 F. Supp. 2d 598, 611 (S.D. Tex. 2010).


\textsuperscript{98} Oracle Brief at 10.
were incurred because of Rimini Street’s purportedly sanctionable misconduct.99

It may be, as Oracle suggests, that the trial court did not invoke its inherent power to award a monetary discovery sanction because it concluded that there was a “statute . . . that adequately addresses the conduct” under Ninth Circuit precedent: Section 505.100 Perhaps “costs may well have been awardeable below as a sanction,” as Oracle argues, even if they were not awardeable as “full costs” under Section 505.101 Because courts inherently possess the power to sanction, the concerns Oracle raises about chilling otherwise meritorious copyright litigation because of the potential for “irretrievably sunk” litigation costs is likely exaggerated.102 Even if it is not, however, the solution to this problem is not to contort the meaning of a federal statute in order to ratify a potentially acceptable (or perhaps implausible but desirable) outcome on unacceptable grounds. “If a statute needs repair, there’s a constitutionally prescribed way to do it. It’s called legislation.”103

VI. Conclusion

As Rimini Street illustrates well, plain meaning can be a trap for the unwary. Dictionaries alone cannot always decide questions of statutory interpretation, even under the plain meaning rule. Construing text requires reading it in its full context, taking into consideration any relevant historical or jurisprudential glosses to the text. In this case, both sides invoke fundamental principles of statutory interpretation. But, reading law here in holistic fashion, Rimini Street makes better sense of Section 505’s “full costs” when viewed in its full context.

99 Id. at 53.

100 Id.

101 Oracle Opposition at 23.

102 Oracle Brief at 49–51.

John Marshall’s Jurisprudence Supports Preemption of California’s Net Neutrality Law

By Randolph J. May & Seth L. Cooper

Telecommunications & Electronic Media Practice Group

About the Authors:

Randolph J. May is President and Seth L. Cooper is a Senior Fellow of the Free State Foundation, an independent, non-partisan free market-oriented think tank located in Rockville, Maryland. This article is adapted, with substantial revisions, from a “Perspectives from FSF Scholars” paper published by the authors on November 2, 2018, on the Free State Foundation website.

Note from the Editor:

The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the authors. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:


It may be hard to see a connection between steamboats plowing the waterways of our early republic and today’s high-speed broadband networks carrying the bits and bytes of internet transmissions. But there is a jurisprudential connection between Chief Justice John Marshall’s 1824 decision in *Gibbons v. Ogden* and the Federal Communications Commission’s (FCC or Commission) 2018 assertion of authority to preempt state laws interfering with interstate internet traffic. In *Gibbons*, Marshall established federal supremacy under the Constitution’s Commerce Clause to preempt a New York law that interfered with steamboat traffic between New York and New Jersey. Marshall determined that the New York law conflicted with a congressional act licensing coastal steamboat traffic, and that it therefore could not be enforced. As Marshall famously put it: “Congress may control state laws so far as it may be necessary to control them for the regulation of commerce.”

*Gibbons* often is considered one of Chief Justice Marshall’s three most important opinions. So it’s worth considering the relevance of the Great Expounder’s *Gibbons* opinion even to a matter as utterly contemporary, and as important to interstate commerce, as today’s internet. First, we will examine the FCC’s January 2018 *Restoring Internet Freedom Order*—in which it asserted preemptive authority to invalidate state laws in conflict with the agency’s declared internet policy—and California’s reaction to the order. Then, we will show how the foundation laid in *Gibbons*, where Marshall was faced with incompatible federal and state laws, buttresses the current FCC’s authority to keep the internet free from conflicting state regulation.

The *Restoring Internet Freedom Order* (RIF Order) repealed the public utility-like regulations the Obama Administration FCC imposed on broadband internet service providers (ISPs) in March 2015. The repealed 2015 regulations became known as the FCC’s *Title II Order*, and they included bright-line bans on blocking, throttling, and paid prioritization, as well as a vague, open-ended “general conduct” standard barring unreasonable interference with end users’ access to internet services or disadvantaging of content providers. These provisions are popularly referred to as

1 22 U.S. 1 (1824).
2 Id. at 206.
3 The other two are, of course, *Marbury v. Madison*, 5 U.S. 137 (1803), and *McCulloch v. Maryland*, 17 U.S. 316 (1819).
5 See *Protecting and Promoting the Open Internet*, WC Docket No. 14-28, Report and Order on Remand, Declaratory Ruling, and Order, 30 FCC Rcd 5601 (2015) (hereinafter Title II Order) (The 2015 order is often referred to as the “Title II Order” because, as explained below, the FCC classified ISPs as common carriers under Title II of the Communications Act in order to impose the public utility-like regulations that are
“net neutrality” regulations. The RIF Order also repealed the Title II Order’s assertion of FCC authority to review internet network interconnection agreements.

The RIF Order reclassified broadband internet access services as Title I “information services” rather than Title II “telecommunications services,” the classification that had been adopted in the 2015 Title II Order. An abundance of federal court and agency precedents treat information services as inherently interstate—therefore within the federal government’s power to regulate under the Constitution’s Commerce Clause—and as non-regulated, or at most lightly regulated, services. Thus, in its 2018 RIF Order, the Commission said, “it is well-settled that Internet access is a jurisdictionally interstate service because ‘a substantial portion of Internet traffic involves accessing interstate or foreign websites.’” Further, “it is impossible or impractical for ISPs to distinguish between intrastate and interstate communications over the Internet or to apply different rules in each circumstance.”

The FCC emphasized in the RIF Order that it was acting consistently with Congress’ established policy in Section 230(b) of the Telecommunications Act of 1996 “to preserve the vibrant and competitive free market that presently exists for the Internet . . . unfettered by Federal or State regulation.” The Commission declared it was returning to “a calibrated federal regulatory regime based on the pro-competitive, deregulatory goals of the 1996 Act.”

Because some states had already voiced their opposition to the FCC’s proposed repeal of the 2015 regulations, the RIF Order directly addressed the legal implications of its deregulatory policy for state and local regulation: “We therefore preempt any state or local measures that would effectively impose rules or requirements for any aspect of broadband service that we address in this order.” In support of the RIF Order’s preemptive authority, the Commission relied on agency precedent recognizing that “federal preemption is preeminent in the area of information services.”

The Commission also relied on modern federal preemption jurisprudence, arguing that “[f]ederal courts have uniformly held that an affirmative federal policy of deregulation is entitled to the same preemptive effect as a federal policy of regulation.”

Despite the FCC’s assertion of preemptive authority in the RIF Order, several states have considered regulating ISPs more stringently than the FCC, and a few have actually done so. Perhaps not surprisingly, given that Silicon Valley web giants like Google and Facebook support net neutrality regulation, California has adopted the most far-reaching state law so far, and the one most unreservedly in conflict with the FCC’s deregulatory policy. Whether California’s law—and other state laws that may come in its wake—survives constitutional scrutiny depends on whether the FCC is right that its deregulatory RIF Order has preemptive effect and is therefore the supreme law of the land in the field of ISP regulation.

This article will examine the FCC’s assertion of preemptive authority in light of the new California law. And it will do so in the context of examining anew Chief Justice Marshall’s Commerce Clause jurisprudence, primarily Gibbons v. Ogden. Many articles review the myriad judicial decisions on preemption in the context of various federal-state conflicts, including federal-state conflicts arising from FCC actions. In many of these “conflict preemption” cases, the Constitution’s Commerce Clause undergirds and supports the assertion of federal authority, but it goes unmentioned. Although it is often taken for granted by courts and commentators, Marshall’s Commerce Clause jurisprudence is the foundation for the exercise of much preemption authority, and it is certainly pertinent to an examination of the lawfulness of the FCC’s assertion of preemption authority in the RIF Order.

I. California Senate Bill 822 and the Department of Justice’s Lawsuit

On September 30, 2018, California Governor Jerry Brown signed SB-822 into law. SB-822 is an attempt to reimpose, at the state level, many of the same restrictions contained in the now repealed Title II Order. SB-822 categorically bans ISPs from blocking access to lawful websites, “throttling” or impairing service, or implementing “paid prioritization” opportunities. SB-822 also includes a provision that closely resembles the RIF Order’s vague “general conduct” standard by prohibiting ISPs from unreasonably interfering with or disadvantaging the communications of customers or competitors. The California law also asserts regulatory authority over “ISP traffic exchange,” a form of regulation of interconnection among ISPs. The RIF Order, by contrast, expressly disclaims authority to regulate interconnection.

Additionally, in at least two significant respects, SB-822’s restrictions are even more stringent and far-reaching than those contained in the Title II Order. First, the law bars mobile broadband service providers from offering California consumers...
“free data” plans that allow consumers to access content from selected websites without such access counting against their monthly data allotments. Second, SB-822 appears to restrict broadband service providers from offering so-called “non-broadband Internet access data services” or “specialized services” over the same last-mile facilities over which they offer broadband internet access services. These services were permitted by the RIF Order.

As soon as Governor Brown signed SB-822, the U.S. Department of Justice (DOJ) filed a lawsuit against California in the U.S. District Court for the Eastern District of California. Subsequently, several ISPs filed another lawsuit challenging SB-822 in the same district. DOJ’s lawsuit against California seeks a federal court order declaring SB-822’s restrictions on broadband internet access services preempted and therefore invalid. In its complaint, DOJ alleges that “SB-822 conflicts with the 2018 Order’s affirmative federal ‘deregulatory policy’ and ‘deregulatory approach’ to Internet regulation” that was adopted in furtherance of Congress’ policy to preserve a competitive free market for the internet “unfettered by Federal or State regulation.” DOJ’s complaint also alleges that SB-822 in the same district.

Given the likelihood that the California law would not survive judicial review, it is not surprising that California agreed to defer its implementation pending judicial review of the RIF Order. Nevertheless, if California ever decides to try to implement its law, DOJ’s lawsuit should succeed on the merits because it is solidly based on modern federal preemption jurisprudence. For instance, the RIF Order cited Arkansas Electric Cooperative Corporation v. Arkansas Public Services Commission, in which the Supreme Court declared that “a federal decision to forgo regulation in a given area may imply an authoritative federal determination that the area is best left unregulated, and in that event would have as much preemptive force as a decision to regulate.” It also cited Minnesota Public Utilities Commission v. FCC, where the U.S. Court of Appeals for the Eighth Circuit stated that “deregulation” is a “valid federal interest[] the FCC may protect through preemption of state regulation.” Similarly, the Eighth Circuit ruled in September 2018 that Minnesota’s attempt to regulate an ISP’s interconnected Voice over Internet Protocol service was preempted because it “attempted regulation of an information service [that] conflicts with the federal policy of nonregulation.”

While these modern federal preemption precedents would likely suffice to assure DOJ of victory in its challenge to SB-822, Chief Justice Marshall’s decision in Gibbons v. Ogden provides further support and shows that this result is deeply rooted in American jurisprudence. Preemption of state laws that affect interstate commerce—like California’s—is by no means novel. Marshall’s early interpretation of the Commerce Clause’s reach supports the preemptive effect of the RIF Order in the following respects:

- In Gibbons, Marshall declared that Congress’ power under the Commerce Clause “applied to all the external concerns of the nation, and to those internal concerns which affect the states generally.” And Marshall recognized that interstate and intrastate services may be “intermingled” in a way which “cannot stop at the external boundary line of each state.” Marshall held that, where such intermingling obtains, Congress has power to regulate despite the presence of some intrastate features of the commerce in question. The RIF Order argued that it is “well settled that Internet access is a jurisdictionally interstate service” because a substantial portion of internet traffic accesses interstate or foreign

---

16 See Complaint, American Cable Assoc., et al., v. Becerra, Case No. 18-01552 (E.D. Cal.) (filed Oct. 3, 2018).
17 DOJ Complaint, at ¶ 41.
18 Id. at ¶ 42.
19 See Mozilla v. FCC, Case Nos. 18-1051, et al. (D.C. Cir, filed Feb. 22, 2018). Petitioners seeking review of the FCC’s RIF Order in the D.C. Circuit case include edge providers like Mozilla and Etsy and public interest groups like Public Knowledge.
21 461 U.S. at 383 (cited by RIF Order, at ¶ 194 n.726).
22 483 F.3d at 580-581 (cited by RIF Order, at ¶ 194 n.726).
24 Several states are attempting to resurrect net neutrality prohibitions by purporting to use their procurement authority to require that ISPs offering proprietary services to the state adhere to prohibitions like those repealed by the FCC’s RIF Order. These actions relating to proprietary procurement service offerings raise somewhat different issues and are not the subject of this article. For our views on this subject, see Seth L. Cooper, State Executive Orders Reimposing Net Neutrality Regulations Are Preempted by the Restoring Internet Freedom Order, Perspectives from FSF Scholars, February 2, 2018, http://www.freestatefoundation.org/images/State_Executive_Orders_Reimposing_Net_Neutrality_Regulations_Are_Preempted_by_the_Restoring_Internet_Freedom_Order.pdf.
25 Gibbons, 22 U.S. at 195.
26 Id. at 194.
generally.” And he stated that “the word ‘among’ means regulate commerce “applied to all the external concerns of the cannot stop at the external boundary line of each State, but intermingled with,” and that thus “commerce among the States

• In Gibbons, Marshall stated that “the acts of New York must yield to the law of Congress” when they “come into collision.” Similarly, in the RIF Order, the FCC determined that laws like California’s which impose net neutrality mandates that the FCC has repealed are inconsistent with the federal deregulatory policy for internet services. In other words, they are, as Marshall put it, in “collision” with the federal policy and must yield.

• In Gibbons, Marshall defined “the power to regulate” as the power “to prescribe the rule by which commerce is to be governed.” The RIF Order prescribes what the FCC variously describes as a “light touch” or “deregulatory” approach for broadband internet services as the general rule. In other words, what the FCC announced as “the federal deregulatory policy restored in this [RIF] order,” consistent with Gibbons, is the rule by which internet commerce is to be conducted.

II. The RIF ORDER AFFIRMED THAT INTERNET ACCESS SERVICES ARE JURISDICTIONALLY INTERSTATE SERVICES THAT CANNOT BE SEGREGATED FROM ANY INTRASTATE ELEMENTS

Chief Justice Marshall’s opinion in Gibbons is foundational to understanding the FCC’s assertion of preemption authority in its RIF Order. Gibbons concerned the lawfulness of New York’s grant of an exclusive operating license to a steamboat company. The exclusive licensing regime impeded steamboat commerce between points in New York and New Jersey because it limited the number of steamboat companies allowed to operate there. DOJ’s challenge to SB-822 concerns state restrictions that the FCC claims impede commerce—not by steamboats, but by digital communications streaming between and among the states and foreign countries.

In Gibbons, Marshall explained that Congress’ power to regulate commerce “applied to all the external concerns of the nation, and to those internal concerns which affect the states generally.” And he stated that “[t]he word ‘among’ means intermingled with,” and that thus “[c]ommerce among the States cannot stop at the external boundary line of each State, but may be introduced into the interior.” The FCC’s conclusions regarding the interstate nature of internet access services in the RIF Order are consonant with Marshall’s exposition of Congress’ power to regulate commerce in Gibbons. In the RIF Order, the Commission, citing several precedents, concluded that it is well settled that internet access is a jurisdictionally interstate service because “a substantial portion of Internet traffic involves accessing interstate or foreign websites.” The agency also stated that “the record continues to show that broadband Internet access service is predominantly interstate because a substantial amount of Internet traffic begins and ends across state lines.” And the Commission determined that state laws like California’s that impose stringent net neutrality mandates that the Commission has repealed “could pose an obstacle to or place an undue burden on the provision of broadband Internet access service and conflict with the deregulatory approach we adopt today.”

The Commission went on to argue that the intrastate and interstate elements of internet access services are so intermingled that Congress has power over the internet access services as a whole; this conclusion is consonant with Marshall’s exposition of Congress’ power to regulate commerce in Gibbons. In the RIF Order, the Commission concluded that, because of the way that modern digital networks route internet traffic, “it is impossible or impracticable for ISPs to distinguish between intrastate and interstate communications over the Internet or to apply different rules in each circumstance.” This is consistent with the recognition by courts that the Commerce Clause allows Congress to preempt state regulation of “activities that inherently require a uniform system of regulation” and those that “impair the free flow of materials and products across state borders.”

Both interstate and intrastate communications include substantial portions of internet traffic that access interstate and foreign websites. Likewise, today’s broadband internet networks transmit data among and within the borders of different states and across the globe. Even an internet communication that begins and ends in the same state may be comprised of bits that traverse many states and even countries overseas before being reconnected to complete the transmission. In other words, the intrastate and interstate elements of broadband internet services are indeed “intermingled” in a way that it makes impossible or impractical to segregate them in a way that they “stop at the external boundary line of each state.” As the Commission concluded, “any effort

27 RIF Order, at ¶ 199.
29 RIF Order, at ¶ 194-195.
30 Gibbons, 22 U.S. at 196.
31 RIF Order, at ¶ 196.
32 Gibbons, 22 U.S. at 195. See U.S. Const. art. I, sec. 8. Constitutional historian Maurice Baxter observed, “The part of the opinion that was the most impressive at the time and would be most durable in the future was a comprehensive exegesis of the commerce clause.” Maurice G. Baxter, The Steamboat Monopoly: Gibbons v. Ogden, 1824 48 (1972).
33 Id. at 194.
34 RIF Order, at ¶ 199 (citing Bell Atl. Tel. Cos. v. FCC, 206 F.3d 1, 5 (D.C. Cir. 2000); NARUC Broadband Data Order, 25 FCC Red 5051, 5024 n.24 (2010); High-Cost Universal Service Support Order, 24 FCC Red 6475, 6496 n.69 (2008)).
35 Id.
36 Id. at ¶ 195.
37 Id. at ¶ 200.
38 Nat’l Ass’n of Optometrists & Opticians v. Harris, 682 F.3d 1144, 1154-55 (9th Cir. 2012).
39 Gibbons, 22 U.S. at 194.
III. The California Law Conflicts with Federal Broadband Internet Policy and Is Therefore Preempted

Chief Justice Marshall observed in Gibbons that the Constitution’s framers foresaw occasions when a state law would come into conflict with a law passed by Congress, and that they provided for such occasions by including the Supremacy Clause in the Constitution. The Supremacy Clause is found in Article VI, Section 2, and it states:

This Constitution, and the Laws of the United States which shall be made in Pursuance thereof . . . shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the Contrary notwithstanding.41

Marshall explained in Gibbons that the Supremacy Clause applies to “such acts of the State Legislatures as do not transcend their powers, but, though enacted in the execution of acknowledged State powers, interfere with, or are contrary to, the laws of Congress made in pursuance of the Constitution or some treaty made under the authority of the United States.”42 “In every such case,” concluded Marshall, “the act of Congress or the treaty is supreme, and the law of the State, though enacted in the exercise of powers not controverted, must yield to it.”43

Three years after his Gibbons opinion, Justice Marshall had occasion in Brown v. Maryland to once again discuss the Commerce Clause, this time with reference to a Maryland law that required importers of foreign goods to pay a fee to obtain a license to sell their products in Maryland. Marshall found that Maryland’s licensing regime that allowed the state to decide what goods could be imported into the state, subject to imposition of importation fees, conflicted with a federal law generally authorizing the importation and sale of goods.44 Marshall declared the Maryland law invalid because federal law is supreme in the event of a conflict with a state law on a matter impacting interstate or foreign commerce:

It has been observed that the powers remaining with the states may be so exercised as to come in conflict with those vested in Congress. When this happens, that which is not supreme must yield to that which is supreme. This great and universal truth is inseparable from the nature of things, and the Constitution has applied it to the often interfering powers of the general and state governments, as a vital principle of perpetual operation.45

The FCC’s RIF Order expressly “preempt[s] any state or local measures that would effectively impose rules or requirements that we have repealed or decided to refrain from imposing . . . or that would impose more stringent requirements for any aspect of broadband service.”46 The RIF Order makes clear that broadband internet service should be governed “by a uniform set of federal regulations, rather than by a patchwork that includes separate state and local requirements.”47 But California’s law reimposes at the state level many of the same restrictions on ISPs contained in the 2015 Title II Order that the FCC repealed in the RIF Order. Indeed, it adopts net neutrality requirements that are even more stringent than those in the 2015 order. Thus, SB-822 clearly conflicts with the RIF Order and the articulated congressional policy “to preserve the vibrant and competitive free market that presently exists for the Internet and other interactive computer services, unfettered by federal and state regulation.”48 Consistent with Marshall’s understanding of the Supremacy and Commerce Clauses, California’s law should be preempted.

IV. The RIF Order’s Deregulatory Rule for Internet Services Is Consistent with Marshall’s Commerce Power Rule

In Gibbons, Chief Justice Marshall defined “the power to regulate” commerce among the states as the power “to prescribe the rule by which commerce is to be conducted.”49 Consistent with Marshall’s view in Gibbons, the FCC’s affirmative decision to adopt a deregulatory approach is a rule by which commerce is to be conducted, in this case a rule by which internet access services will be regulated by the federal government under the Commerce Clause power.

In the RIF Order, the Commission declared that it was adopting “a calibrated federal regulatory regime based on the pro-competitive, deregulatory goals of the 1996 Act.”50 In the Commission’s view, an affirmative decision to adopt a deregulatory rule is still an exercise of its power to regulate under the Commerce Clause. Contrary to claims by some advocates of regulation, 51

---

40 RIF Order, at ¶ 200.
41 The proposal to add the Supremacy Clause to the Constitution was adopted after the delegates rejected Madison’s idea of allowing a federal veto of any state law. Much later, in 1833, Madison wrote in a letter to future president John Tyler: “The necessity of some constitutional and effective provision guarding the Constitution and the laws of the union against violations of them by the laws of the states was felt and taken for granted by all, from commencement to the conclusion of the work performed by the convention.” See James Madison to John Tyler (1833) (unsent), 3 Max Farrand, Records of the Federal Convention 527 (Reprint 1996).
42 Gibbons, 22 U.S. at 211.
43 Id.
44 25 U.S. 419 (1827).
45 Id. at 448.
46 RIF Order, at ¶ 195.
47 Id. at ¶ 194.
49 Gibbons, 22 U.S. at 196.
50 RIF Order, at ¶ 194.
51 See, e.g., Barbara van Schewick, Geo. Brown Signs SB 822, Restoring Net Neutrality to California, The Center for Internet and Society Blog (September 30, 2018) (arguing “the FCC cannot prevent the states from adopting net neutrality protections because the FCC’s repeal order removed its authority to adopt such protections”), https://cyberlaw.stanford.edu/blog/2018/09/george-brown-signs-sb-822-restoring-net-neutrality-california; Comments of Public Knowledge, Implementation of Section 621(a)(1) of the Cable Communications Policy Act of 1984 as Amended by the Cable Television Consumer Protection and
the Commission did not simply abandon authority in this area and leave matters up to the states. Rather, the Commission's reestablishment of what it referred to as “an affirmative federal policy of deregulation” was a deliberate exercise of regulatory power consistent with Marshall's understanding of the term.\(^{52}\)

The D.C. Circuit has recognized that “providing interstate [communications] users with the benefit of a free market and free choice” is a “valid goal” and that “[t]he FCC may preempt state regulation . . . to the extent that such regulation negates the federal policy of ensuring a competitive market.”\(^{53}\) The Commission's establishment of a carefully calibrated federal regulatory regime, albeit a light-touch rather than a heavy-handed one, is a rule establishing a federal policy under the Commerce Clause—and one that supports preemption of conflicting state laws.

V. Conclusion

If ultimately litigated to its conclusion, DOJ's lawsuit challenging California's SB-822 likely will succeed based on modern federal preemption precedents. But it is important to understand that Chief Justice Marshall's jurisprudence, especially his opinion in the landmark *Gibbons v. Ogden* case, supplies a critical constitutional backdrop for those modern precedents. Consideration of Marshall's Commerce Clause jurisprudence deepens and reinforces the conclusion that the federal deregulatory policy applicable to broadband internet access services reestablished in the FCC's *RIF Order* should, and most likely will, result in the preemption of California's net neutrality law and any similar laws that might be passed in other states.

---

\(^{52}\) *RIF Order*, at ¶ 194.

Concealed Carry and the Right to Bear Arms

By Joseph G.S. Greenlee

Civil Rights Practice Group

About the Author:

Joseph Greenlee is an attorney in Steamboat Springs, Colorado, a fellow in constitutional studies and firearms policy at the Millennial Policy Center, and a policy advisor for legal affairs at The Heartland Institute.

Note from the Editor:

The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:


The Second Amendment protects “the right of the people to keep and bear Arms.” In recent times, what it means to bear arms has become the subject of some debate. That bearing arms involves the public carrying of arms to some extent is clear enough, but to whom the right extends, where it extends, and in what manner remains unsettled.

This article addresses what manner of carrying the Second Amendment protects—specifically, whether the concealed carrying of arms is protected. The Supreme Court, American history and tradition, and the most influential lower court decisions indicate that it is.

I. Heller

The Supreme Court expressly defined “bear arms” in District of Columbia v. Heller. Adopting a definition Justice Ruth Bader Ginsburg had previously provided, the Court determined that the “natural meaning of ‘bear arms’” is to “wear, bear, or carry . . . upon the person or in the clothing or in a pocket, for the purpose . . . of being armed and ready for offensive or defensive action in a case of conflict with another person.” Carrying “in the clothing or in a pocket” is concealed carry, whereas wearing “upon the person” includes open carry. Thus, the Supreme Court explicitly included both concealed carry and open carry in its definition of “bear arms.”

The Court did note, however, that “the Second Amendment is not unlimited” and recognized that historically “the right was not a right to carry . . . in any manner whatsoever.” Rather, states have been permitted to regulate the manner of carrying. As the Court pointed out, “the majority of the 19th-century courts to consider the question held that prohibitions on carrying concealed weapons were lawful.” The Court cited cases that upheld such bans when open carry remained available. Thus, the Supreme

1 U.S. Const. amend. II. In full, the Second Amendment provides: “A well-regulated Militia, being necessary to the security of a free State, the right of the people to keep and bear Arms, shall not be infringed.”

2 See, e.g., James Bishop, Hidden or on the Hip: The Right(s) to Carry After Heller, 97 Cornell L. Rev. 907, 922 (2012) (“Under any application of originalist analysis . . . states may not prohibit open carry unless they instead offer the alternative outlet of concealed carry.”); Saul Cornell, The Right to Carry Firearms Outside of the Home: Separating Historical Myth from Historical Realities, 39 Fordham Urb. L.J. 1695, 1696 (2012) (“Apart from service in militia, there is little evidence of a broad constitutional consensus on a right to carry arms in public.”); Jonathan Meltzer, Open Carry for All: Heller and Our Nineteenth-Century Second Amendment, 123 Yale L.J. 1486, 1528 (2014) (“[T]he Second Amendment protects the right to carry openly outside the home.”).


4 Id. at 584 (quoting Muscarello v. United States, 524 U.S. 125, 143 (1998) (Ginsburg, J., dissenting)).

5 Id. at 626.

6 Id.
Court explained that the right to bear arms includes concealed carry and open carry, but it suggested that a state can regulate the manner of carrying—for instance, by prohibiting concealed carry if open carry is available.

Given this, it is constitutional for a state to prohibit open carry while broadly allowing concealed carry—as some states do today? The “original meaning” sources relied on by the Heller Court, the right-to-carry cases extolled by the Heller Court, and post-Heller decisions from lower courts indicate that the right to bear arms is not infringed as long as law-abiding citizens are able to publicly bear arms either openly or concealed.

II. The Founding Era

The Heller Court focused on the founding-era understanding of the right to bear arms. To that end, it found Noah Webster’s definitions of “keep,” “bear,” “arms,” and “militia” persuasive. While the Court had no need in Heller to provide the entire definition of “bear,” it is worthy of closer examination here.

Webster’s definitions of “bear” included: “To wear; to bear as a mark of authority or distinction; as, to bear a sword, a badge, a name; to bear arms in a coat.” This authoritative source expressly contemplated bearing arms as carrying a concealed firearm.

Moreover, Webster defined “pistol” as “A small fire-arm,” and he explained in his definition that “Small pistols are carried in the pocket.” Notably, as Webster explained in defining “gun,” pistols were never called guns in the founding era. “Gun” referred to a long gun. With the understanding that pistols were regularly carried in a concealed manner, the framers could have codified the right to bear “guns” rather than “arms” had they intended to exclude concealed carry. Or they could have expressly excluded it as some state constitutions later did. But they did neither, nor did they ever demonstrate an intention of excluding concealed carry from the Second Amendment’s protections in any other way. In fact, pistols, knives, swords, and armor were ubiquitous militia equipment throughout the colonial and founding eras and included in Webster’s definition of “arms,” demonstrating that the Second Amendment was intended to protect much more than just long guns.

Concealable firearms in America date back to the first permanent English settlement. In 1622, “300 short pistols with fire locks” were delivered to Jamestown Colony. Indeed, it was common practice in the founding era to carry concealed firearms. Historian George C. Neumann explained that “[a]mong eighteenth-century civilians who traveled or lived in large cities, pistols were common weapons. Usually they were made to fit into pockets.” Similarly, in describing founding-era America to his friend in Scotland in 1775, a Virginian wrote, “No person goes abroad without his sword, or gun, or pistols.”

As indicated by Webster’s definitions, pistols were commonly carried in one’s pocket. Consequently, a popular pistol size was referred to as “pocket pistols.” “Pocket pistols, also known as coat pistols, were small in size yet of large caliber that could easily be carried in one’s trouser pocket or, more commonly, the coat pocket.” Larger versions were referred to as “overcoat pistols.” A smaller size was called “muff pistols,” because women would commonly conceal them in their hand warmer muffets. Muff pistols “were quite popular . . . in the 18th century” and included Queen Anne pistols. “The Queen Anne style of pistol first became popular in England during the reign of Queen Anne (1702-1714).” Their popularity soon spread throughout the colonies, and the pistols were later used by soldiers in the French and Indian War and in the American Revolution. Other firearms designed to be concealed were “boot pistols,” which “could be easily concealed high in the top of riding boots.”

Many other pistols existed in the colonial and founding eras, and with one exception they were never prohibited from being carried either openly or concealed. The exception was a 1686 New Jersey law that prohibited concealed carry by anyone, as well as the open carrying of swords, pistols, and daggers by planters. Planters were “those who settled new and uncultivated

---

7 Id. at 581, 582, 584, 595.
15 3 American Archives, 4th series, 621 (Peter Force ed., 1840) (Sept. 1, 1775).
16 Jeff Kinard, Pistols: An Illustrated History of their Impact 57 (2003). Derringers, which entered the market in the 1820s, became the most popular and well-known pocket pistols. Id.
18 Id.
21 Id. at 56. A popular variation of the boot pistol was the underhammer pistol, invented in the first half of the nineteenth century. “Such handy weapons were considered indispensable on the frontier and along highways and back alleys of the new nation.” Id. at 57.
22 23 The Grants, Concessions, and Original Constitutions of the Province of New-Jersey 289–90 (1758).
Thus, frontiersmen could openly carry long guns, but not handguns. People in towns could openly carry anything. Significantly, “[n]o colony followed New Jersey’s statute against concealed carry, or the restrictions on open handgun carry by planters. Nor did any state until about half a century after American independence.”

Laws that required colonists to carry arms were more common. Many colonies mandated that colonists bear arms to church, court, musters, or to work on the roads or in the fields. None mandated the manner in which arms were to be carried.

III. The Nineteenth Century

The first states to restrict the bearing of arms were Kentucky and Louisiana, which each banned concealed carry in 1813. Throughout the nineteenth century, other states enacted similar restrictions. Far from coming to a consensus, courts reached a variety of conclusions when the laws were challenged. In his annotations to James Kent’s famous Commentaries on American Law, future Supreme Court Justice Oliver Wendell Holmes, Jr. noted that “it has been a subject of grave discussion, in some of the state courts, whether a statute prohibiting persons, when not on a journey, or as travellers, from wearing or carrying concealed weapons, be constitutional. There has been a great difference of opinion on the question.” The Supreme Court of Georgia exclaimed, “tot homines, quot sententiae.—so many men, so many opinions!”

Of them all, it is most instructive to review the cases the Supreme Court relied on to define the individual right in Heller. In defining the Second Amendment right, the Heller Court approvingly cited five cases interpreting the right to bear arms protected by the Second Amendment or analogous arms-bearing rights in state constitutions.

The 1813 Kentucky ban was ruled unconstitutional in Bliss v. Commonwealth, where the Court of Appeals of Kentucky held that a prohibition on either concealed or open carry would violate the right to bear arms. Conversely, the 1813 Louisiana ban was upheld by the Supreme Court of Louisiana in State v. Chandler, where the court stated that open carry was the guaranteed right.

The Alabama Supreme Court upheld a concealed carry ban in State v. Reid in 1840, declaring that the legislature had “the right to enact laws in regard to the manner in which arms shall be borne . . . as may be dictated by the safety of the people and the advancement of public morals.” The manner selected would be valid as long as the arms could still be used for self-defense efficiently:

We do not desire to be understood as maintaining, that in regulating the manner of bearing arms, the authority of the Legislature has no other limit than its own discretion. A statute which, under the pretence of regulating, amounts to a destruction of the right, or which requires arms to be so borne as to render them wholly useless for the purpose of defence, would be clearly unconstitutional. But a law which is intended merely to promote personal security, and to put down lawless aggression and violence, and to that end inhibits the wearing of certain weapons, in such a manner as is calculated to exert an unhappy influence upon the moral feelings of the wearer, by making him less regardful of the personal security of others, does not come in collision with the constitution.

In other words, a state may regulate the manner in which arms can be carried if it promotes public safety and still allows the carrier to defend herself.

A few years later, in Nunn v. State, the Georgia Supreme Court followed Reid’s reasoning in upholding a prohibition on concealed carry while striking a restriction on open carry. This...
holding may seem to indicate that open carry is constitutionally protected and concealed carry is not. But it is more plausible that the court required that one or the other be available, and that its holding was intended to reflect the legislature’s preference for open carry—which it demonstrated by prohibiting concealed carry while merely regulating open carry. Regarding the concealed carry ban, the court said, “it is valid, inasmuch as it does not deprive the citizen of his natural right of self-defense.”38 Since open carry was available, citizens could still defend themselves. But had open carry been prohibited also, the concealed carry ban would have deprived citizens of the natural right of self-defense and therefore would have violated the Second Amendment.

Similarly, after creating some uncertainty earlier in that century, the Tennessee Supreme Court held in Andrews v. State that a general carry “prohibition is too broad,” but “[i]f the Legislature think proper, they may by a proper law regulate the carrying of this weapon publicly, or abroad, in such a manner as may be deemed most conducive to the public peace, and the protection and safety of the community from lawless violence.”39

Of these cases, only the Chandler case indicated that concealed carry was not protected by the right to bear arms, declaring that open carry “is the right guaranteed by the Constitution of the United States.”40 But even Chandler was later interpreted by the Louisiana Supreme Court as “prohibiting only a particular mode of bearing arms which is found dangerous to the peace of society.”41 Based on changes in societal preferences, a present-day law regulating open carry but allowing concealed carry could arguably serve the same purpose.42

IV. Post-\textit{Heller}

Since \textit{Heller}, many courts have decided whether the right to bear arms includes concealed carry. Like the pre-\textit{Heller} cases, there is a difference of opinion among various courts; also like the pre-\textit{Heller} cases, they generally agree that the right protects both open and concealed carry.

In challenges to concealed carry permitting schemes, the First,43 Second,44 Third,45 and Fourth46 Circuits all assumed (without deciding) that concealed carry is protected. The D.C. Circuit and Seventh Circuit have gone further, both determining that concealed carry is protected.

When the District of Columbia banned open carry and required applicants to show a “good reason” for needing a concealed carry permit, the D.C. Circuit struck it down in \textit{Wrenn v. D.C.} because the burden on concealed carry was too great.47 The court held that restrictions on the manner of bearing arms were permissible, but that “the law must leave responsible, law-abiding citizens some reasonable means of exercising” the right.48 Thus, a “shall-issue” scheme was required, where permits are generally issued to all applicants who meet objective criteria.49

The Seventh Circuit struck down Illinois’ complete prohibition on bearing arms in \textit{Moore v. Madigan}. Illinois responded by enacting a shall-issue licensing scheme for concealed carry. This scheme was upheld in 2016, indicating that a prohibition on open carry and a shall-issue licensing scheme for concealed carry was consistent with the Seventh Circuit’s understanding of the Second Amendment.50

In \textit{Norman v. State}, the Florida Supreme Court followed the approach of the D.C. and Seventh Circuits—as well as the overall theme of the cases summarized in this article—in rejecting a challenge to Florida’s open carry ban.51 The court determined that the state’s shall-issue licensing scheme satisfied the constitutional requirement because it “provides almost every individual the ability to carry a concealed weapon.”52 Since anyone not prohibited by law from owning a gun could carry one concealed, the state could regulate the open carrying of arms.

By contrast, only the Ninth and Tenth Circuits have upheld concealed carry bans. But both did so without considering the availability of open carry. In the Tenth Circuit case, \textit{Peterson v. Martinez}, the plaintiff “repeatedly expressed . . . that he is not challenging the Denver ordinance” restricting open carry, so the court conducted its analysis “based on the effects of the state statute [restricting concealed carry] rather than the combined effects of the statute and the ordinance.”53

The Ninth Circuit took a similar approach in \textit{Peruta v. County of San Diego}, although in that case the court took it upon itself to consider only concealed carry rather than the combined effects of the laws prohibiting all carrying.54 Subsequently, in last year’s \textit{Young v. Hawaii}, a three-judge panel of the Ninth Circuit decided that since concealed carry is unavailable, open carry must be permitted.55 The court is currently considering whether to rehear that case en banc, along with another case that challenges open carry and concealed carry bans simultaneously. The latter

---

38 Id. at 251 (emphasis omitted).
42 See infra section V.
43 Gould v. Morgan, 907 F.3d 659 (1st Cir. 2018).
44 Kachalsky v. Cty. of Westchester, 701 F.3d 81 (2d Cir. 2012).
45 Drake v. Filko, 724 F.3d 426 (3d Cir. 2013).
46 Woollard v. Gallagher, 712 F.3d 865 (4th Cir. 2013).
47 864 F.3d 650 (D.C. Cir. 2017).
48 Id. at 663.
51 215 So.3d 18 (Fla. 2017).
52 Id. at 28.
53 707 F.3d 1197, 1208 (10th Cir. 2013).
54 824 F.3d 919 (9th Cir. 2016).
55 896 F.3d 1044 (9th Cir. 2018).
The case, *Flanagan v. Becerra*, was filed in response to *Peruta*. It challenges the combined effects of California’s open and concealed carry restrictions to ensure that the court considers the full context of the burden on the right to bear arms—thus precluding the possibility of the court considering either restriction in a vacuum as it did in *Peruta*.

V. PUBLIC POLICY

If the right to bear arms does not protect concealed carry at all, it would follow that it protects only open carry. To be sure, one can still exercise the core right of self-defense with an openly carried firearm. But most Americans prefer concealed carry. There are roughly 17.25 million concealed carry permitholders in America, and this does not account for concealed carriers in the fourteen states that do not require a permit. Many millions of these Americans would not carry at all if they had to carry openly. As UCLA law professor Adam Winkler explained, “for those who want fewer guns on the streets, there are a million reasons to prefer open carry,” including that “[v]ery few gun owners want to carry openly displayed guns.”

If concealed carry were held not to be part of the right to bear arms at all, it could become far less available. States compelled to allow open carry would be less inclined to allocate the funds and resources necessary to administer a concealed carry licensing scheme. For instance, they may instead license open carry. And anti-gun states that currently view concealed carry as the lesser evil may abolish their concealed carry schemes since open carry would be permissible either way.

Regardless of prospective policy considerations, American history and tradition show that the carrying of concealed arms is part of the right protected by the Second Amendment. It can be prohibited only if open carry is available, just as open carry can be prohibited only if concealed carry is available.

VI. CONCLUSION

The Supreme Court has elucidated that the scope of the Second Amendment is defined by the founding-era understanding of the right, as informed by American history and tradition. A historical analysis shows that both concealed and open carry are protected by the right, and that a government may only restrict one if the other remains available for law-abiding citizens to exercise.


In his new book, *51 Imperfect Solutions: States and the Making of American Constitutional Law*, Sixth Circuit Judge Jeffrey Sutton advocates for a renewed focus on state constitutional law. American constitutional law is dominated by court decisions—both state and federal—interpreting the federal constitution. The “critical conviction” of Judge Sutton’s book is that “a chronic underappreciation of state constitutional law” has distorted the shape of state and federal law and skewed “the proper balance between state and federal courts in protecting individual liberty.” Too many issues have been nationalized, in Judge Sutton’s view, because courts have resolved challenges to a state’s action under the federal constitution without first considering what that state’s constitution has to say about the matter. Judge Sutton suggests that this tendency has diminished respect for state constitutional guarantees and trust in state court judges. Judge Sutton’s book articulates a distinctly federalist view of constitutional law, and it is filled with ideas that conservatives and liberals alike will both applaud and question.

I. Judge Sutton’s Four Examples Illustrating the Interaction Between State and Federal Courts

The heart of the book is four stories about the interaction between state and federal courts over whether and how much to protect four specific individual rights. Judge Sutton uses these stories to make his case for putting the states at the “vanguard” of American constitutional law. Each story describes the complex interactions between state and federal courts as they decide which constitutional rights to recognize, with the state courts portrayed as the heroes of each episode.

The most powerful story is about how state and federal courts responded to the eugenics movement of the early 20th century. Many readers will remember from law school Justice Oliver Wendell Holmes’ infamous line that “three generations of imbeciles are enough,” which concluded his opinion in *Buck v. Bell*, an 8-1 Supreme Court decision approving the forced sterilization of a mentally disabled person. Fewer, however, know that several state courts had held similar forced sterilization laws unconstitutional before that 1927 decision. Those state court decisions almost uniformly held that state eugenics laws violated due process or equal protection guarantees. But after *Buck v. Bell*, most state courts “fell in line” with the Supreme Court’s decision for decades after *Buck v. Bell*. These courts acted as though
a decision under the federal constitution preempted the field of what state constitutions might have to say about eugenics laws.

Judge Sutton identifies several lessons from this episode in American constitutional law, two of which are particularly notable. The first, a theme that runs throughout his book, is that when the Supreme Court definitively resolves a complicated national policy debate as a matter of federal constitutional law, its judgments can stifle further constitutional debate at the state level—even when, as a legal matter, states remain free to act as they please. Judge Sutton’s second lesson is the flipside of the first: state courts, he says, should resist the urge to interpret their own constitutions “in reflexive imitation” of federal law. As he puts it elsewhere in his book, relying too heavily on the Supreme Court as the “guardian of our rights” runs “the risk of creating state courts that lack the necessary fortitude to fill the gaps when we need it most.” That is an important lesson, but it invites further questions about an equally important responsibility of courts at all levels to recognize that courts and constitutions need not resolve all policy debates, and for jurists to have the humility and fortitude to leave some issues to the legislatures and the people.

Another chapter uses the exclusionary rule to tell a slightly different story of how state courts approach constitutional rights before and after landmark Supreme Court decisions. Judge Sutton traces the development of the exclusionary rule from the early state court cases rejecting it through the Supreme Court’s adoption of an exclusionary rule for federal prosecutions in its 1914 decision in Weeks and its 1920 decision in Silverthorne Lumber. At the time, the Supreme Court left the states free to decide whether to exclude illegally obtained evidence in state prosecutions. States responded in a variety of ways. Some slowly adopted an exclusionary rule through legislation, others embraced exclusionary principles through court decisions interpreting their state constitutions, and, as Judge Sutton recounts, many rejected the exclusionary rule altogether. That is, until the Supreme Court’s 1961 decision in Mapp v. Ohio nationalized the exclusionary rule.

Judge Sutton views the development of the exclusionary rule as “a story with some potentially promising features.” Chief among them, from his perspective, is that the Supreme Court did not impose a nationwide exclusionary rule right away. This allowed the states “to decide for themselves how to weigh the costs and benefits of evidentiary exclusion.” And this state-level experimentation “provided more empirical information about the pros and cons of exclusion” before settling on a single constitutional rule for state and federal prosecutions. Judge Sutton points out that the states took seriously their responsibility to decide for themselves whether their constitutions required an exclusionary rule. Unlike in the aftermath of Buck v. Bell, the state courts did not automatically adopt the federal rule. Having an example of state courts continuing to grapple with the meaning of their own constitution even after the Supreme Court has weighed in is essential in a book whose central thesis is that state courts should do so more often.

51 Imperfect Solutions then picks up the development of the exclusionary rule with the Supreme Court’s 1984 decision in United States v. Leon, which adopted a “good faith” exception to Mapp’s exclusionary rule. Judge Sutton recounts the states’ surprising reaction to that decision: twenty states rejected Leon through court decisions construing their own constitutions. Even as Judge Sutton praises the state courts’ independence, he worries that some courts made their decisions for the wrong reasons, elevating policy concerns over a careful interpretation of the text and traditions underlying their state constitutional guarantees. Others might worry that the state court reaction to Leon casts doubt on whether state courts can ever truly escape the shadow of landmark Supreme Court decisions on federal constitutional questions. State courts, which were the original skeptics of the exclusionary rule, became its biggest defenders after having the rule forced upon them, even when the Supreme Court later gave them greater flexibility. In that respect, the story of how state courts dealt with the exclusionary rule looks remarkably like the story of the state court response to eugenics laws.

The third story in 51 Imperfect Solutions concerns court challenges to school board policies that did not exempt Jehovah’s Witnesses from requirements to salute the flag and participate in the pledge of allegiance. This episode is less about the relationship between state and federal constitutional guarantees, and more about how public opinion can influence court decisions at all levels. Judge Sutton recounts that state and federal courts were initially unwilling to take the Jehovah’s Witnesses’ claims seriously. That changed after the Supreme Court decided Minersville School District v. Gobitis in 1940, rejecting a free exercise challenge to a local policy requiring participation in a flag-salute ceremony. Within four years, two state supreme courts had interpreted their own constitutions to provide the protections Gobitis denied, and the U.S. Supreme Court largely reversed course in West Virginia Board of Education v. Barnette, where it found that forced participation in flag-salute ceremonies amounted to compelled speech forbidden by the First Amendment. Why such a rapid change? Public reaction to Gobitis was swift and overwhelmingly negative. Judge Sutton catalogs 170 newspaper editorials criticizing the Gobitis decision; The New Republic, for example, published an editorial drawing parallels to Nazi Germany. The state courts may have been “path blazers” in the sense that they published their decisions retreating from Gobitis first, but Judge Sutton’s account suggests that it was the American people—not state courts—who moved the path of American constitutional law in this instance.

In his last example, school funding, Judge Sutton tells a different type of story. This story starts with the Supreme Court’s 1973 decision in San Antonio Independent School District v. Rodriguez rejecting a federal constitutional right to equal funding among public school districts. In response to that decision, litigants raised state constitutional challenges to their systems of funding public schools; Judge Sutton counts victories for those litigants in twenty-seven states. Judge Sutton identifies two potential reasons for reformers’ greater success at the state level. He points out that state courts setting rules for one state and one school system face fewer institutional challenges than a federal court seeking a single national rule for fifty diverse states and systems. And state courts are interpreting state constitutions, many of which have language directly addressing public education, which is often favorable for school funding reform advocates. Whereas the federal constitution largely places limits—rather than duties—on the federal government, Judge
Sutton points out that many state constitutions also “impose obligations on government.”

Along the way, Judge Sutton asks whether education funding advocates actually benefited from losing in Rodriguez. He doubts the federal courts would have been willing to go as far as state courts have in protecting school funding, calling this a “federalism discount” baked into federal constitutional rulings. And he questions whether state courts would have been as receptive to state constitutional challenges if the U.S. Supreme Court had already recognized some version of a federal constitutional right to adequate funding. State courts, Judge Sutton argues, are more receptive to constitutional claims when there are clear lines of accountability placing the burden to protect individual rights on state governments. Otherwise, Judge Sutton worries that state courts may use federal constitutional guarantees as an excuse to do nothing.

II. THE VIRTUES AND VICES OF JUDGE SUTTON’S STATE-FIRST APPROACH TO CONSTITUTIONAL DECISIONMAKING

This is one of many distinctively federalist arguments Judge Sutton makes throughout 51 Imperfect Solutions. The central theme of Judge Sutton’s book is that courts and litigants too often overlook the fact that our individual liberties do not flow exclusively from the federal constitution. Our federal system provides two layers of protection for individual liberties—one at the federal level and another at the state level through state constitutions and other state laws. State constitutions thus play an important role in protecting liberty. In his book, Judge Sutton contends the legal system should take those state constitutional guarantees more seriously by treating them as the separate and independent barriers protecting individual liberty that the founders envisioned.

Judge Sutton’s book also embraces federalism in other ways. He envisions a state-first approach to recognizing constitutional rights, where the states are the first ones to decide whether to recognize a constitutional right, and where they do so as a matter of state constitutional law. Only after the state courts have weighed in would the federal courts decide whether to adopt a uniform constitutional rule that applies nationwide. As Judge Sutton notes, this allows state courts to adopt different constitutional rules that respect and honor differences among the states. The language, history, and tradition underlying state constitutional protections differ, so it would be surprising if every state adopted the same answer to a particular constitutional question. And if state courts err in their interpretation, state constitutions are far easier to amend than the federal constitution. For some constitutional questions, Judge Sutton hopes that the states may arrive at a range of acceptable solutions tailored to local circumstances that eliminates the need for a uniform federal rule.

While many will find Judge Sutton’s commitment to federalism attractive, his state-first approach to constitutional decisionmaking is likely to be controversial. There is no obvious way to implement his proposal, which depends on federal courts at least temporarily abstaining from resolving federal constitutional challenges in favor of letting state courts decide similar issues on state constitutional grounds first. Although the U.S. Supreme Court can avoid resolving constitutional questions while they percolate in the state courts, other courts generally cannot. Federal courts of appeals, and even many state courts themselves, do not have similar flexibility (or at least do not have that flexibility if they are not prepared to recognize the asserted right as a matter of state law). They generally must resolve the constitutional questions presented to them by litigants. Few litigants will willingly shelve a federal constitutional claim, and in many cases may actively prefer one, precisely to obtain a decision with nationwide consequences.

Judge Sutton also advocates for a more modest version of his state-first approach. Under this model, state courts would always resolve state constitutional challenges before turning to parallel federal claims—even if it means articulating why the state constitution rejects a right that the federal constitution clearly recognizes. The courts of only three states—Oregon, Maine, and New Hampshire—have adopted this model, and few courts are likely to follow suit. Busy judges rarely look for more work, especially when addressing an issue will not affect the outcome of the case. But judicial economy is not the only objection; there is also wisdom in the principle that courts should not issue advisory opinions. The quality of judicial reasoning and analysis generally drops as courts stray from the issues that actually matter to the outcome of a case, reflecting natural human tendencies given limited time and attention spans. It is thus far from clear that this state-first model would increase the attention paid to state constitutional interpretation or the quality of state constitutional decisions, however laudable those goals potentially are in the abstract.

A more fundamental problem lies with the premise of 51 Imperfect Solutions, which “take[s] for granted that vigorous individual rights protection by some court is beneficial” while recognizing “that may not always be the case.” In other words, the book offers a take about which courts (state or federal) interpreting which constitution (state or federal) would be better suited to recognize new constitutional rights protecting individual liberties. This, of course, assumes that some court should be recognizing a particular right in the first place.

Indeed, a legitimate critique of 51 Imperfect Solutions is that some courts might interpret it as a call for more judicially recognized constitutional rights. One could read the first 175 pages of the book as a call for common-law constitutionalism, where the existence and scope of constitutional rights are developed largely through a back-and-forth conversation between state and federal courts. That dialogue inevitably tilts towards constitutionalizing rights at some level—state or federal—rather than considering the role of the legislatures in protecting rights by statute. Moreover, because incorporation ensures that almost all federal constitutional protections automatically apply to the states as well, state courts may have little to add to the conversation other than extending constitutional protections beyond those federal law already affords.

But reading Judge Sutton’s book as a call for judicial activism would be a mistake. Despite saying little about how courts should interpret constitutional provisions, 51 Imperfect Solutions subtly argues that state courts should interpret their constitutions based on text and tradition. Judge Sutton suggests that, when state courts recognize rights that federal courts have not (or symbolically reject rights that federal courts have recognized), they should do so
only through “marshaling the distinct state [constitutional] texts and histories and drawing their own conclusions from them.” If those “first principles” cannot justify recognizing or extending a constitutional right, Judge Sutton suggests state courts should not do so. This is a message many judicial conservatives will applaud, along with Judge Sutton’s rebuke of Justice Brennan’s view a generation ago that (in Judge Sutton’s words) “[s]o long as there is a progressive will . . . there is a new way for granting relief” federal courts denied by imposing the same obligations via creative interpretations of state constitutions.

Progressives and conservatives alike have also paid too little attention to a second, equally important aspect of Judge Sutton’s argument: In his view, federal courts should exercise more judicial restraint. Although Judge Sutton believes state courts should do more to protect individual liberty through their constitutions, he believes federal courts should respond by doing less. The U.S. Constitution “was not designed to facilitate rights innovation,” Judge Sutton argues. The founders “thought of the States as the first bulwarks of freedom,” and Judge Sutton urges his fellow federal judges to allow the state courts to exercise that responsibility by not rushing to nationalize every issue.

III. Key Takeaways

Many will find this a refreshing touch of judicial humility. 51 Imperfect Solutions suggests the “federal-first” approach—treating the federal constitution as providing a national answer to every policy dispute—has slowly eroded trust in the federal judiciary. Many court observers agree. Judge Sutton hopes that re-establishing “[s]tate primacy in guarding individual rights” will restore confidence in both the state and federal judiciaries.

Whether or not Judge Sutton is right about that, 51 Imperfect Solutions itself presents an imperfect solution for restoring that balance. Asking litigants to disarm or for courts to effectively abstain from deciding issues under the federal constitution is asking a lot. Few litigants are interested in partial victories, allowing Utah and California to afford different protections for what they view as fundamental rights. Convincing state court judges to interpret their constitutions based on “local language, context, and history” rather than “[tak[ing] sides on the federal debates and federal authorities” will be a challenging task, as will filling the state bars with advocates who will mine the historical record and present the state courts with those arguments. And without a cultural change, adopting Judge Sutton’s proposals could simply transform each state constitution into a one-way ratchet of ever-expanding rights, which he predicts is “destined to fail over the long term.”

Adopting Judge Sutton’s dual vision—with more active state court judges focused on state-level constitutional sources and more restrained federal judges—would thus require a fundamental shift in the way litigants, courts, and scholars approach constitutional law at all levels. Such a shift would probably have to start with education about our system of federalism, not just in law schools, but in high schools across the country. It is precisely the boldness of what Judge Sutton is really proposing that makes it so thought-provoking. 51 Imperfect Solutions invites a conversation worth having.
The Tenacity of Transformation Theory, and Why Constitutional History Deserves Better

By Stephen B. Presser

Federalism & Separation of Powers Practice Group

A Review of:
The Second Creation: Fixing the American Constitution in the Founding Era, by Jonathan Gienapp
http://www.hup.harvard.edu/catalog.php?isbn=9780674185043

About the Author:
Stephen B. Presser is the Raoul Berger Professor of Legal History Emeritus at Northwestern University's Pritzker School of Law. He is a Visiting Scholar of Conservative Thought and Policy at the University of Colorado, Boulder for 2018-2019.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:

For many years, American progressive scholars have been inclined to view legal and constitutional change as a series of bold “transformations” featuring a switch from one set of values and principles to another. Roscoe Pound (a progressive in his youth, but a conservative at the end of his career) set the template with his Formative Era of American Law, in which he described how creative state and federal judges such as Joseph Story and Lemuel Shaw altered the common law of England to fit the needs of the United States. In 1977, Morton Horwitz, in his aptly titled Transformation of American Law, took Pound’s template and explained that the same judges Pound praised had actually engaged in a pernicious shifting of the legal rules to favor an emerging merchant and entrepreneurial class who had enlisted the newly professional lawyer cohort in their nefarious enterprise.

The intellectual effort to chronicle law as transformation continued with Stanford historian Jack Rakove’s celebrated book, Original Meanings: Politics and Ideas in the Making of the Constitution, which explained how the early framers saw the Constitution as malleable and argued that they understood the Constitution as a loose template to be adjusted as the needs of the nation changed. A purported alteration in James Madison’s thinking about the nature of law and political institutions was a key subject for Rakove. Another book, Mary Sarah Bilder’s Madison’s Hand: Revising the Constitutional Convention, has argued for a similar change in Madison’s beliefs by tracing the evolution of his revisions of his Notes on the Constitutional Convention.

Jonathan Gienapp, author of The Second Creation: Fixing the American Constitution in the Founding Era, is a junior colleague of Rakove’s at Stanford. Like Rakove and Bilder, he considers Madison central to his exposition. But rather than make Madison’s Notes on the Constitutional Convention his subject, Professor Gienapp focuses on the activities of the first Congress, where Madison was a prime mover. Gienapp argues that the first Congress, in essence, engaged in a “second creation” of the Constitution, abandoning one closer to the British “constitution” of broad principles which guaranteed flexibility and change in favor of one fixed in meaning for all time, the interpretation of which relied on an “original understanding” of the document’s framers. In other words, Gienapp sees the first Congress as a moment of constitutional transformation.

The progressive trope of transformation, which implicates the notion of a “living” or “evolving” Constitution, seems like it

1 Roscoe Pound, The Formative Era of American Law (1938)
4 Mary Sarah Bilder, Madison’s Hand: Revising the Constitutional Convention (2015).
was designed, consciously or unconsciously, to support modern judicial progressives, such as the members of the Warren Court or Justices Anthony Kennedy and Sandra Day O’Connor, who understand their task to be to refashion constitutional principles—particularly the “equal protection” and “due process” provisions of that charter—to fit the “evolving standards of decency” that supposedly characterize American civilization.6

Rakove’s, Bilder’s, and Gienapp’s books could be seen as an attack on the jurisprudence of originalists like Justices Antonin Scalia and Clarence Thomas, who embrace the notion that the only sensible and valid strategy of constitutional hermeneutics is to interpret the document according to its plain meaning at the time it was passed or amended. The theory of an evolutionary development of constitutional meaning, based as it is on an idea similar to Darwin’s speculation with regard to the evolution of the species, has undeniable intuitive appeal. Nevertheless, evolutionary jurisprudence is in uneasy tension with more basic ideas about ours being a government of laws and not of men, and thus with our hallowed concept of the rule of law itself. If judges become legislators, there is an end to separation of powers, and popular sovereignty also goes by the board. Some scholars are fighting a rearguard action.7 But Gienapp’s new book and the honors bestowed on previous books telling a similar story—Horwitz’s and Bilder’s books both won the Bancroft Prize, the highest accolade the history fraternity can bestow—show that alternative stories about constitutional and legal development are out of favor.

Horwitz, Rakove, Bilder, and Gienapp are all learned and energetic scholars, and all demonstrate originality, dedication, an appreciation of both the primary sources and the secondary literature. Gienapp also demonstrates some flair with language and a deep and realistic appreciation of the manner in which constitutional theory can be enlisted in the service of particular political ends, or perhaps the manner in which political actors can shade the truth to meet partisan policy needs. His book is a very valuable review of the struggles in the first Congress between the


6 This notion of using the past to justify radical transformation in the present and future was a principal feature of the notable late seventies’ jurisprudence in American law schools, Critical Legal Studies. See, e.g., Debra Livingston, Note, Round and Round the Bramble Bush: From Legal Realism to Critical Legal Scholarship, 95 Harv. L. Rev. 1669, 1678 (1982) (“By demonstrating that first principles, not only doctrinal details, are products of historical circumstance and historically specific modes of legal reasoning, the critical legal scholar uses history to disclose that the underlying assumptions of doctrinal fields lack the necessity sometimes claimed for them—to demonstrate that such assumptions represent mere choices of one set of values over another.”). On Critical Legal Studies generally, see, e.g., Stephen B. Presser, Law Professors: Three Centuries of Shaping American Law, Chapter 14 (2017).


8 For a notable account of the political and judicial struggles in the early republic, see, e.g., GORDON S. WOOD, EMPIRE OF LIBERTY: A HISTORY OF THE EARLY REPUBLIC 1789-1815 (2009).

9 I do not mean to disparage Madison or his contributions. He was a great practitioner of what the framers understood as the “new science of politics,” and his contributions to the Federalist regarding the separation of powers and federalism deserve their immortal fame.


12 For a discussion of the early federal court cases, see, e.g., Stephen B. Presser, The Original Misunderstanding: The English, The Americans, and the Dialectic of Federalist Jurisprudence (1991). In that book I argued, curiously, that there was a transformation of legal thought in the early republic, but the transformation I proposed was quite different from that set forth by Gienapp. I argued that the early American judges and justices were applying a static and religious view of law, while John Marshall laid the groundwork for a flexible, accommodating, and dynamic view of law. See also GARY L. McDOWELL, THE LANGUAGE OF THE LAW AND THE FOUNDATIONS OF AMERICAN CONSTITUTIONALISM (2010) (a brilliant exposition—subtle than mine—making the case that the initial conception of the Constitution as fixed
Nevertheless, there is something about Madison that continues to fascinate and inspire provocative scholarship such as this study, and his position in Congress may be the reason for the focus on the legislative branch. Whatever it means for how we should understand the Constitution, it is clear that Madison's beliefs shifted in the 1790s, and that he therefore moved from an alliance with Alexander Hamilton regarding commerce and a strong central government towards a closer embrace of Jefferson's agrarian and states' rights beliefs. This shift, along with his undeniable importance during the framing and ratification of the Constitution, makes Madison a subject of perpetual interest.13

In the end, though, the transformation rhetoric is not wholly satisfying. One can argue, as Gary McDowell convincingly did,14 that the meaning of the Constitution was fixed at the time of its ratification, and that this was the orthodox and original understanding and indeed the very purpose of creating the document. This is not to deny that our constitutional debates have always featured one group or another challenging this original understanding and arguing for a living Constitution in order to accomplish its own goals, or to transform our politics. Perhaps our constitutional history inevitably involves an ongoing and repetitive series of efforts to replace the original understanding with new constitutional models. In that case, perhaps the appropriate metaphor for American constitutional hermeneutics is a cyclical or repetitive unfolding, rather than a series of transformations.

The distinguished political scientist Garrett Ward Sheldon, in a review of Professor Bilder's book, made an observation which applies just as nicely to Professor Gienapp's notion that the Constitution began as malleable, became temporarily fixed for political purposes, and could therefore morph further in service of other ideologies:

So, this author's perspective is that the underlying philosophies of the U.S. Constitution in Locke's Natural Rights ideology; Montesquieu's separation of powers; Aristotle and Cicero's views of man's social nature and resultant democracy; the Reformed Christian suspicion of human nature and evil (all prevalent in the Founding

period and known by the classically educated Founders), do not provide a stable, solid, permanent foundation for the Constitution and Republic, but we are all subordinate to the vagaries of contemporary political controversies, personalities and interests; relativistic and transitory. And so in this view such mutability of social and government should continue right up to our own time: the Bill of Rights protections of Freedom of Speech, press, religion and Due Process of Law may be modified by future generations, especially if some speech or religion may offend certain groups, or the rights of the accused and judicial procedure are perceived as unfair or irrelevant to certain classes of accusers or aggrieved.15

Just so.

13 On Madison's political thought, see especially, GARRETT WARD SHELDON, THE POLITICAL PHILOSOPHY OF JAMES MADISON (2003). It is important to understand that Sheldon brilliantly traces a consistency in Madison's thought, which led Madison, at various times and as circumstances dictated, to be for or against a strong central government. Thus, Sheldon explains, "Madison's political philosophy historically shifted between Lockean liberal and classical republican, federal and states' rights perspectives, with a consistent view to a balanced, moderate government that accurately reflected the Christian view of human nature as egoistical and domineering, realistically establishing a stable and just regime." Id. at xiv. In Sheldon's view, unlike Gienapp's, no shifting sense of constitutional meaning is necessary to explain or justify Madison's actions. Rather, the explanation is his consistent religious and political philosophy as applied to different contexts.

14 See McDowell, supra note 12.

Can a New Establishment Clause Jurisprudence Succeed in Protecting Religious Minorities Where Lemon Has Failed?

By Alexandra M. McPhee

Religious Liberties Practice Group

About the Author:
Alexandra M. McPhee is an attorney and the Director of Religious Freedom Advocacy at Family Research Council. In her position, she advocates for a robust role for religion in the public square, the advancement of religious liberty in the United States, and an understanding of the Establishment Clause consistent with its original meaning.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:

Samantha Elauf was the face of the 2015 United States Supreme Court decision EEOC v. Abercrombie & Fitch, which vindicated her right under Title VII to be free from religious discrimination in hiring because of her faith-based decision to wear a headscarf. In 1990, the Supreme Court decided that the Free Exercise Clause does not require exemptions from generally applicable laws in Employment Division v. Smith; this decision, which denied Alfred Smith the right to use a controlled substance in a religious ceremony, resulted in the bipartisan passage of the Religious Freedom Restoration Act to protect religious actions like Smith’s at the national level. More than a century earlier, in 1860, Rabbi Morris Raphall was the first rabbi to deliver a prayer opening Congress’ legislative session. “[P]iously bedecked in a white tallit and a large velvet skullcap,” he invoked the blessing of “Lord God of Abraham, of Isaac, and of Jacob,” thanked God for “establish[ing] a Commonwealth after a model of . . . the tribes of Israel, in their best and purest days,” and gave a traditional blessing in Hebrew. What do a Muslim millennial teenage girl, a middle-aged Klamath Native American man, and a nineteenth century rabbi have in common? They each exemplify the accommodation and acceptance of religious minorities in America under the law and in our nation’s history.

In American Legion v. American Humanist Association, the Court is reviewing the U.S. Court of Appeals for the Fourth Circuit’s holding that it is an Establishment Clause violation for a Maryland bi-county commission to own and maintain a cross-shaped veterans’ memorial in Bladensburg, Maryland. Relying on the analytical framework set forth in Lemon v. Kurtzman, the panel majority concluded that because the memorial is forty feet tall, located in a high-traffic intersection, maintained with government funds, and in the shape of a cross, the memorial “has the primary effect of endorsing religion and excessively entangles the government in religion.”

6 American Humanist Ass’n, 874 F.3d at 200.
Several religious minority groups have filed amicus briefs arguing both for and against the constitutionality of the Peace Cross, as it is known to locals. Some of these groups argue that regardless of whether the Fourth Circuit’s decision is affirmed or reversed, the Supreme Court should maintain the current state of the law surrounding the Establishment Clause—especially the Lemon test and its variants—because it either adequately or best protects religious minorities and fosters a pluralistic society. They express concern that a different approach—particularly the “coercion” test advocated by American Legion—would enable majority suppression of minority religious exercise.

This article refutes the claim that current Establishment Clause jurisprudence best protects minority religious groups. It first argues that analysis based on Lemon and later decisions modifying it does not satisfactorily protect minority religions, much less best protect them. These tests are fundamentally flawed because they permit, and even require, subjective judicial decision-making. Next, the article argues that an approach rooted in the original meaning of the First Amendment best protects minority religions. Such an approach provides an objective measure for gauging Establishment Clause violations, in contrast to the subjective reasoning required by the Lemon test and its successors.

Finally, the article argues that reliance on Lemon-based precedent to protect religious minorities is misplaced. The political branches, including local, state, and federal legislative and executive bodies, are better suited than the courts to protect religious minorities and include them in American civic life.

Elauf, Smith, and Rabbi Raphall were able to engage and flourish in the public square consistent with their minority religious beliefs, and they did not need the Lemon test to do so; indeed, judicial intervention in Smith’s case worked against his freedom to practice his religion. The Peace Cross, a veterans’ memorial that invokes imagery from the majority religion of Christianity, does not violate the Establishment Clause and does not harm religious minorities. An originalist interpretation of the Establishment Clause makes this clear, and our country can best determine how to adapt to its increasingly religiously diverse population through conversation and compromise in the political branches.

I. Religious Minority Displays and Practices Are Vulnerable Under Current Law

Many years before courts began interpreting the Establishment Clause, Alexander Hamilton expressed his thoughts on the interpretation of the Constitution in a letter to George Washington. He wrote, “whatever may have been the intention of the framers of a constitution or of a law, that intention is to be sought for in the instrument itself, according to the usual and established rules of construction.”11 Interpreting a constitution or a law is easy when the text of the instrument is clear and its application is uncontroversial. The task is more difficult, however, when the text of a given provision is ambiguous or its application to a case is not obvious.

Many cases resolved under the Establishment Clause fit this description of “difficult,” to say the least. Consider the text of the Clause: “Congress shall make no law respecting an establishment of religion . . . .” The word “Congress” is likely well-known to most readers. So is the word “law.” But what is “an establishment”? Some even argue that the word “religion” is a term of art rather than a reference to religion generally.14 In an ironic twist of jurisprudence, the clearest parts of the clause, “Congress” and “law,” were read out of it in Everson v. Board of Education and several Establishment Clause cases involving government action generally.15 In Everson, the Supreme Court

[Notes and references included in the text]
incorporated the Establishment Clause through the Fourteenth Amendment and applied it to the states.16 The decisions involving government action showed that the Supreme Court was concerned about more than just laws establishing religion. As a result, no government body may engage in any act “respecting an establishment of religion.”

As for that phrase, “respecting an establishment of religion,” there “are only so many lights to assist the courts in arriving with more accuracy at the true interpretation of the intention.”17 Thomas Jefferson offered counsel on how to approach such questions: “On every question of construction,” return “to the time when the Constitution was adopted, recollect the spirit manifested in the debates, and instead of trying what meaning may be squeezed out of the text, or invented against it, conform to the probable one in which it was passed.”18 Jefferson and Hamilton's advice underlie the originalist approach to interpreting constitutional provisions.

But several of the groups opposing the Maryland bi-county commission's ownership and maintenance of the Peace Cross—and even some that favor it—insist that the Court should stand by precedents that neglect the meaning of the words of the Establishment Clause at “the time when [it] was adopted.”19 The Lemon test and its progeny not only depart from the original meaning of the constitutional text they claim to interpret and apply, they also fail to protect religious minority displays and practices as well as a more constitutionally-rooted test would.20

A. The Lemon Test Is Too Subjective

In 1971, the Supreme Court in Lemon v. Kurtzman articulated a three-prong test to determine if a statute passed muster under the Establishment Clause. The Court said that judges should ask whether there was a secular purpose for the statute, whether its primary effect advanced or inhibited religion, and whether it “foster[ed] an excessive government entanglement with religion.”21 In Larson v. Valente,22 the Court “indicate[d] that laws discriminating among religions are subject to strict scrutiny,” but that “laws ‘affording a uniform benefit to all religions’ should be analyzed under Lemon.”23

The Lemon prongs received additional gloss in later decisions. In County of Allegheny v. ACLU Greater Pittsburgh Chapter, a majority of the Supreme Court adopted the “endorsement” test, which asks judges to discern “what viewers may fairly understand to be the purpose of the display” being challenged as an establishment of religion.24 The endorsement test also instructs that “[e]very government practice must be judged in its unique circumstances.”25 The Court would later attribute to the hypothetical viewer, sometimes called “the reasonable person,”26 knowledge of the purpose of the challenged government action.27

In a later case, the Court said that the excessive-entanglement prong of Lemon should be treated “as an aspect of the inquiry into [an action’s] effect” due to their similar analyses.28 Finally, in Van Orden v. Perry, Justice Stephen Breyer, in a concurring opinion, emphasized a judge’s need to use his best “legal judgment” in deciding cases involving religious displays.29 In analyzing the Ten Commandments monument at issue, he considered how the display was “used” and “the context of the display,” including the message conveyed, “the physical setting,” and the period of time over which the display went unchallenged.30 Justice Breyer’s opinion essentially represented a return to the endorsement test.

Throughout all of these decisions, the Court progressively moved from calling Lemon’s three prongs “must have[a]” to calling them “no more than helpful signposts” or “familiar considerations.”31 But these tests are still in force and are used together or separately depending on the given government action and the circuit court of appeals rendering the decision. The problem is that these tests are malleable, such that judges at the trial level must make decisions without clear guidance on the extent to which they should rely on or extrapolate from the

---

20  Sutherland, supra note 17, § 363, at 696.
21  Lemon, 403 U.S. at 612-13 (italics added).
25  County of Allegheny, 492 U.S. at 595 (quoting Lynch, 465 U.S. at 694 (O’Connor, J., concurring)).
27  McCrery Cty. v. ACLU, 545 U.S. 844, 862 (2005).
30  Id. at 700-03.
31  Lemon, 403 U.S. at 612; Hunt v. McNair, 413 U.S. 734, 741 (1973); McCreary Cty., 545 U.S. at 859.
evidence before them. And the outcome of an appealed decision is still unpredictable, as circuit court judges review lower court decisions de novo because those decisions involve mixed questions of law and fact. Interpreting and applying the Establishment Clause should not be this complicated. Indeed, according to one treatise author, this jurisprudence imperils the separation of powers. Rules of construction “are a part of the law of the land equally with the statutes themselves, and not much less important. The function of such interpretation unrestrained by settled rules would introduce great uncertainty, and would involve a power virtually legislative.” Unfortunately, when it comes to interpreting the Establishment Clause jurisprudence has indeed “introduce[d] great uncertainty” and has become “interpretation unrestrained” and “a power virtually legislative.”

B. Current Law Surrounding Religious Minority Displays and Practices Lacks Consistent, Principled Reasoning

An examination of relevant caselaw demonstrates that neither the Lemon test, the endorsement test, nor legal judgment have provided an effective shield for religious minorities. The following survey of decisions shows that, while religious minorities sometimes successfully use Lemon and its successors to combat constitutional violations, there is no guarantee that they will succeed nor a consistent standard to predict what will happen when a government tries to accommodate their displays or practices.

1. Holiday Displays

In County of Allegheny, the Justices’ analyses of the constitutionality of displaying a menorah on public property splintered in multiple directions. Justice Harry Blackmun would have upheld the menorah based on his idiosyncratic belief that the menorah was secular enough to be constitutionally displayed on public property. Justice Sandra Day O’Connor concluded that the menorah, while a religious symbol, passed constitutional muster because the county situated it next to a holiday-themed tree. Three Justices considered the menorah a religious symbol and would have held its presence on public property unconstitutional. Four other Justices considered the entire display constitutional regardless of the religiosity of the menorah or its surrounding props.

After the Court issued County of Allegheny, a New York district court held that a display of a menorah next to a tree decorated with lights was a religious display that violated the Establishment Clause. The lighted tree, though secular, could not counter the religious significance of the menorah. The court distinguished the display from the one in County of Allegheny on the ground that the tree’s Christmas lights were obscured during the day. The court opined that the reasonable observer would think the city was displaying an eighteen-foot menorah next to a plain old tree and therefore endorsing Judaism.

Skoros v. City of New York and Mehdi v. United States Postal Service reveal how government officials have had to employ policies that afford little room for logic because of the fractured outcome of County of Allegheny. In Skoros, the Second Circuit determined that a public school holiday display policy from the New York State Department of Education did not violate the Establishment Clause. The policy considered a nativity to be a “religious symbol” but a menorah and crescent moon and star to be “secular symbols” for the purposes of classroom holiday displays. The Second Circuit described the policy as a “good-faith—if not entirely correct—reading of the Supreme Court’s decision in Allegheny.” It ultimately concluded that the policy was simply a constitutional means of carrying out the secular purpose of “promot[ing] pluralism through multicultural holiday displays.”

In Mehdi, challengers unsuccessfully argued that the United States Postal Service’s seasonal display policy, which permitted “evergreen trees bearing nonreligious ornaments’ and ‘menorahs (when displayed in conjunction with other seasonal matter),’” violated the Establishment Clause. The challengers argued that the policy failed to include the display of what the challengers characterized as the non-religious crescent moon to represent Muslim practices around the same time of year. In upholding USPS’s seasonal display policy, the district court observed that the “policy was no doubt crafted by the Postal Service with Allegheny in mind.”

32 See, e.g., Skoros v. City of New York, 437 F.3d 1, 26 (2006) (citing Lynch, 465 U.S. at 693-94 (O’Connor, J., concurring)) (“[N]o specific evidence is necessary to allow judges to determine how a mature objective mind would process the images and information conveyed by a holiday display.”). See, e.g., id. at 13 (“Where, as here, a case is tried on a stipulated record, our review is de novo because the district court’s rulings are necessarily conclusions of law or mixed fact and law.”).
33 2 Sutherland, supra note 17, § 363, at 696.
34 Id.
35 County of Allegheny, 492 U.S. at 613-14.
36 Id. at 636, 633 (O’Connor, J., concurring).
37 Id. at 637-38 (Brennan, J., concurring in part and dissenting in part).
38 Id. at 655 (Kennedy, J., concurring in part and dissenting).
39 Id. at 526-27.
40 Skoros, 437 F.3d 1.
42 Skoros, 437 F.3d at 19.
43 Id. at 22.
44 Id.
45 Id.
46 Mehdi, 988 F. Supp. at 729.
47 Id. at 723-24 & n.3.
2. Government Accommodations of Religious Practices

Judicial decisions applying Lemon or its variants may result in favorable outcomes for religious minorities, like cases involving government approval of eruv. But courts’ Lemon-based analyses in such cases lack uniformity, which undermines any notion of the test's stability and fails to guide governments as they make policy.

Eruvs are a ceremonial religious practice of some Orthodox Jewish sects in which adherents put up wires between utility poles to demarcate certain areas where members of the sects live and worship. Typically, members of these sects are prohibited “from pushing or carrying objects outside their homes on the Sabbath or Yom Kippur.” But adherents “may engage in such activities outside their homes on the Sabbath within an eruv.” An eruv “extends the space within which pushing and carrying is permitted on the Sabbath beyond the boundaries of the home, thereby enabling ... [adherents] to push baby strollers and wheelchairs, and carry canes and walkers, when traveling between home and synagogue.” In one case involving an eruv setup, the Third Circuit reasoned that “a reasonable, informed observer would not perceive an endorsement of Orthodox Judaism because the Borough’s decision to approve the eruv would ‘reflect[] nothing more than the governmental obligation of neutrality toward religion.’” The Second Circuit reasoned similarly in a challenge involving an eruv setup in New York.

The Second Circuit went on to argue, however, that the accommodation of eruv had “more of a secular purpose, cause[d] less of an advancement of religion, and foster[ed] less church-and-state entanglement” than allowing “a ‘private Christian organization for children’ to hold meetings at a public school for the purpose of conducting religious instruction and Bible study” or “a Christmas nativity scene display, on public property,” which earlier Supreme Court decisions had upheld. This was supposed to be so because the eruv were not alleged to “contain any overtly religious features that would distinguish them to a casual observer as any different from strips of material that might be attached to utility poles for secular purposes.” This reasoning suggests that some religious practices will pass muster under the Establishment Clause and others will not simply based on whether the average viewer—as imagined by the judge deciding the case—knows their religious significance. Like the Second Circuit, a New Jersey district court noted the eruv’s “almost invisible boundary” and that “[a]n eruv does not in any way force other residents to confront daily images and symbols of another religion.” The district court also noted that “the eruv itself has no religious significance or symbolism and is not part of any religious ritual.” This reasoning harkens back to County of Allegheny’s question of the religiosity of a given symbol or display as the determinative factor in whether a government action violates the Establishment Clause.

These cases show that the Lemon test and those derived from it do not provide the protection for minority religions that some advocates think they do. Courts deciding cases under Lemon have no choice but to perpetuate the absence of a clear rule of law. Their decisions inevitably devolve into statements about their own “legal judgment” or highly fact-specific determinations, neither of which provide a stable basis for governments trying to decide whether they may constitutionally approve or accommodate a given religious practice or display.

II. Religious Minority Displays and Practices Will Be Better Protected by a Clear Legal Standard Rooted in the Original Meaning of the Establishment Clause

The answer to confusion over what constitutes an unconstitutional establishment of religion is an objective test that relies on more than Supreme Court precedent accumulated from 1947 to 2005. Establishment Clause law has suffered from a lack of principled guidelines according to which judges can render decisions. As a result, decisionmakers often “interpolat[e] meaning into a legal text instead of interpreting meaning from the text.” Between 1947 and 2005, only one case, Marsh v. Chambers, articulated an objective standard. To determine whether prayers in Congress violated the Establishment Clause, the Court analyzed historical practices at the time of the founding and the ratification of the First Amendment. The Court has increasingly incorporated this kind of reasoning in its decisions, including in Hosanna-Tabor Evangelical Lutheran Church and School v. EEOC and Town of Greece v. Galloway. This historical approach—or an originalist interpretation of the Constitution’s text—is the right approach to Establishment Clause challenges.

Critics of this approach argue that it would leave the Establishment Clause without teeth. This is likely true to the

50 See Jewish People for the Betterment of Westhampton Beach v. Vill. of Westhampton Beach, 778 F.3d 390, 393 (2d Cir. 2015); Tenafly Eruv Ass’n v. Borough of Tenafly, 309 F.3d 144, 176 (3d Cir. 2002); American Civil Liberties Union v. Long Branch, 670 F. Supp. 1293, 1293 (D.N.J. 1987); Smith v. Cmty. Bd. No. 14, 491 N.Y.S.2d 584, 588 (Sup. Ct. 1985).
51 Tenafly Eruv Ass’n, 309 F.3d at 152 (citation omitted).
52 Id.
53 Id.
54 Id. at 176 (alteration in original).
55 Jewish People for the Betterment of Westhampton Beach, 778 F.3d at 395. See also Smith, 491 N.Y.S.2d at 587 (similar reasoning in a state trial court decision).
56 Jewish People for the Betterment of Westhampton Beach, 778 F.3d at 396 (emphasis in original).
57 Id. (citing Good News Club, 533 U.S. at 103-04; Lynch, 465 U.S. at 671).
58 Id. at 395.
60 Id. at 1296.
62 Tenafly Eruv Ass’n, 778 F.3d at 393 (2d Cir. 2015).
63 Id. at 790 (“[H]istorical evidence sheds light not only on what the draftsmen intended the Establishment Clause to mean, but also on how they thought that Clause applied to the practice authorized by the First Congress—their actions reveal their intent.”).
64 Hosanna-Tabor Evangelical Lutheran Church and School v. EEOC, 565 U.S. 171 (2012).
extent that the original meaning of the Establishment Clause would permit religious displays on public property (and to the extent that the critic wants religious displays to be held unconstitutional). As one court of appeals judge observed, “There is, put simply, lots of history underlying the practice of placing and maintaining crosses on public land . . .” 66 But applying an originalist approach opens up the public square for the expression of all religions. While a court applying a historical interpretation of the Establishment Clause would likely approve a monument like the Peace Cross, it would also likely uphold similarly situated displays inspired by minority religions.

An originalist understanding of the Establishment Clause would make judicial decision-making more objective and stable than it is under current law. As one scholar explains, “This approach requires the judge to look at the text of the Constitution, and if it is unclear, the judge tries to discover not what the text ought to mean but what it did mean to those who wrote the words and, more importantly, to those who voted for those words to become law.” 67

The American Legion, which supports the Peace Cross memorial, argues that a “coercion” test is the best way to implement the original meaning of the Constitution’s prohibition of religious establishments. It would prohibit “government actions that pose a realistic threat to religious liberty—those that coerce belief in, observance of, or financial support for religion.” 68 The Lemon test and the succeeding tests are not viable because they do not “accord[] with history and faithfully reflect[] the understanding of the Founding Fathers,” and because “the text and history of the First Amendment show the Establishment Clause was designed to prohibit coercion.” 69 Religious displays like the Peace Cross, the American Legion argues, should only be found unconstitutional if they are found coercive. 70

One amicus argues against an adoption of this analysis, claiming that “[a] narrower standard that . . . focuses only on coercion would open the door to sectarian endorsements that will aggravate religious tensions and needlessly divide Americans.” 71 He further describes American Legion’s analysis as a “break[] with [the Court’s] Establishment Clause precedents.” 72 It is true that an approach rooted in an originalist interpretation is narrower than the analytical frameworks found in Lemon and its progeny. But an originalist interpretation of the Establishment Clause would set clearer boundaries for which religious displays or practices are acceptable, which would be fairer and more predictable than current law. It is difficult to say exactly how many more religious displays would be considered acceptable under a consistently applied standard based on an originalist interpretation, but recent jurisprudence indicates that principled boundaries would be no less helpful to religious minorities than to members of majority faiths.

In Town of Greece v. Galloway, in which the Court adopted an original understanding of the Establishment Clause with respect to legislative prayer, 73 the Court said it was “virtually inconceivable that the First Congress, having appointed chaplains whose responsibilities prominently included the delivery of prayers at the beginning of each daily session, thought that this practice was inconsistent with the Establishment Clause.” 74 As American society has grown more religiously diverse, figures including the Dalai Lama, Rabbi Joshua Gruenberg, Satguru Bodhinatha Veylanswami, and Imam Nayyar Imam have opened legislative sessions with statements expressly declaring their deeply held religious beliefs. As the Court said, Congress “acknowledges our growing diversity not by proscribing sectarian content but by welcoming ministers of many creeds.” 75

But under Lemon and succeeding tests, courts often proscribe government support of an action or display simply because it is sectarian. A court found a menorah unconstitutionally on public property because its presence next to an unlighted tree in the daytime would appear to the reasonable person to be a government’s endorsement of religion. 76 Even when they uphold religious displays or accommodations, courts employ inconsistent reasoning, 77 which gives no guidance to officials. In that sense, the Lemon decision and its successors render policymakers’ options more narrow because they are forced to make rigid, if not totally nonsensical, distinctions between what displays and practices are “in” or “out.” This is what happened in Skoros and Medhi, 78 where public school and post office officials were forced to write policies based on the outcome of County of Allegheny: menorahs and decorated Christmas trees “in,” nativities and crescent moons and stars “out.”

Judges, public officials, and citizens deserve guidance. American Humanist Association observed during oral argument that these “cases are ill-suited for sweeping pronouncements and categorical rules,” 79 and arbitrary court decisions and government policies show why. Relying on the original meaning of the Establishment Clause is the best way to protect religious minorities because that meaning is fixed. A standard that relies on an originalist interpretation, while not perfect, provides more consistent guidance than a jurisprudence that relies on “what

67 Kelsey, supra note 61, at 21.
68 Opening Br. at 23, American Legion, Nos. 17-1717 & 18-18.
69 Id. at 18 (alterations in original) (citation omitted); id. at 24 (altering capitalization).
70 See id. at 19.
71 Kalsi brief at 3-4.
72 Id. at 10.
73 Town of Greece, 572 U.S. at 602 (adopting the reasoning used in Marsh and stating that the decision “reflected the original understanding of the First Amendment”) (Alito, J., concurring).
74 Id. at 602-03.
75 Id. at 579 (majority opinion).
76 Ritell, 466 F. Supp. 2d at 525.
77 See Jewish People for the Betterment of Westhampton Beach, 778 F.3d at 393; Tenafly Eruv Ass’n, 309 F.3d at 176; Long Branch, 670 F. Supp. at 1293; Smith, 491 N.Y.S.2d at 588.
78 Oral Argument Transcript at 83, American Legion, Nos. 17-1717 & 18-18.
III. Religious Minority Displays and Practices Will Be Better Protected by the Political Branches

Some amici argue that current Establishment Clause jurisprudence better protects minority religions than a more originalist approach would. An originalist approach “would tempt some governments to erect crosses and some citizens to pressure government to do so,” and it would allow a government to “endorse its preferred religious teachings and be candid about what it was doing.” An approach like American Legion’s coercion test, others argue more specifically, would not “address the danger that the majority will, through government endorsements of its own faith, marginalize minority groups” or that “members of the majority [will] claim[] religious superiority, slinging allegations of religious inferiority at minorities.”

These fears are unfounded. Governments are still subject to the “push and pull of the political process—above all from accountability for their speech through the democratic process,” Professor Hillel Y. Levin has argued that “courts are not typically the appropriate forum for delineating the required accommodations” for minority religions. Indeed, “the track record for those who seek religious accommodations in court is not particularly favorable;” they often lose. Furthermore, Professor Michael McConnell states “that the Court’s intervention over the last forty years has made things worse, not better.” In the realm of Establishment Clause law, litigation outcomes for religious accommodations are unpredictable; even where they have won, courts’ reasoning varied such that future outcomes remained uncertain.

The political branches, however, have demonstrated that they can protect religious minority rights and respond to America’s increasingly pluralistic society. Of course, courts have an important role in protecting religious minorities, but as Professor Levin argues, the need for judicial intervention is the exception and not the rule. The political branches have shown themselves to be more efficient sources of great protection for religious minority beliefs and practices. In Samantha Elauf’s case, the Supreme Court agreed with the EEOC that Abercrombie & Fitch violated Title VII of the Civil Rights Act of 1964 when Elauf’s religion became “a motivating factor” in its hiring process because her decision to wear a headscarf conflicted with Abercrombie & Fitch’s “Look Policy” prohibiting “any” head gear. The case, which the Supreme Court decided based on the statute’s text, shows that legislation passed by Congress and enforced by an executive agency can protect minority religious practice. According to the Court, “Congress defined ‘religion,’ for Title VII’s purposes, as includ[ing] all aspects of religious observance and practice, as well as belief.” And

Title VII does not demand mere neutrality with regard to religious practices—that they be treated no worse than other practices. Rather, it gives them favored treatment, affirmatively obligating employers not “to fail or refuse to hire or discharge any individual . . . because of such individual’s ‘religious observance and practice.’”

This case demonstrates the role the courts should play in protecting minority religions: interpreting a law passed by a legislative body and applying it.

Alfred Smith’s case shows how political “institutions often respond to judicial decisions that are unfavorable to religious groups by expanding religious minority groups’ rights.” Smith had lost his job as a counselor because he used a controlled

---

79 County of Allegheny, 492 U.S. at 595 (quoting Lynch, 465 U.S. at 694 (O’Connor, J., concurring)); Van Orden, 545 U.S. at 700.
80 Baptist Joint Committee brief at 36-37.
81 Kalsi brief at 7-8; Muslim Advocates brief at 7.
82 Freedom from Religion Found., Inc. v. City of Warren, 707 F.3d 686, 697 (6th Cir. 2013). There is historical precedent for relying on the political branches to protect religious rights. Thomas Jefferson and the Virginia General Assembly committed the timeless principles announced in Virginia’s Act for Religious Freedom to the legislature:

[W]e well know that this Assembly, elected by the people for the ordinary purposes of legislation only, have no power to restrain the acts of succeeding assemblies constituted with powers equal to our own . . . ; yet we are free to declare, and do declare, that the rights hereby asserted are of the natural rights of mankind; and that if any act shall be hereafter passed to repeal the present, or to narrow its operation, such act will be an infringement of natural right.

84 Id. at 1642.
86 Levin, supra note 83, at 1640-41. In a case like Tenafly Eruv Association v. Borough of Tenafly, a court’s intervention would be welcome. 309 F.3d at 151, 155. In that case, members of an Orthodox Jewish sect wanted to put up eruvos and received permission from the borough to do so. After strong pushback from citizens who did not want the eruvos in place and the discovery of an ordinance that prohibited certain attachments to poles on public land, the borough voted to remove the eruvos previously put up with its approval. Eruv supporters challenged the borough’s vote on the ground that it violated their First Amendment right to free exercise, among other claims. They proved during litigation that the borough did not enforce the ordinance equally, permitting private postings, house number signs, or church direction signs. Id. at 151, 155 (citations omitted). The borough argued that it had a compelling interest to avoid an establishment clause violation. The Third Circuit rejected the borough’s argument, stating that “a reasonable, informed observer would not perceive an endorsement of Orthodox Judaism because the Borough’s change of heart would ‘reflect[] nothing more than the governmental obligation of neutrality toward religion.” Id. at 176 (alteration in original).
87 Abercrombie & Fitch Stores, Inc., 135 S. Ct. at 2031.
88 Id. at 2032-34.
89 Id. at 2033 (alteration in original) (quoting 42 U.S.C. §2000e(j)).
90 Id. at 2033-34 (alteration in original) (emphasis added) (citation omitted).
91 Levin, supra note 83, at 1642.
substance, peyote, for a Native American religious practice. He sought unemployment benefits, but the employment division denied his application because he was terminated for "work-related misconduct."99 The employment division did not exempt him from its policies because his violation took place in the course of his religious exercise, and neither did the Supreme Court. Justice Antonin Scalia, writing for the majority in Employment Division v. Smith, said, "Respondents urge us to hold, quite simply, that when otherwise prohibited conduct is accompanied by religious convictions, not only the convictions but the conduct itself must be free from governmental regulation. We have never held that, and decline to do so now."93 Congress responded to the decision with a concerted, bipartisan effort to protect the religious freedom rights of people like Smith, whose religious practices would clash with the law unless exempted.94 The Religious Freedom Restoration Act, signed into law by President Bill Clinton in 1993,95 was created "in order to provide very broad protection for religious liberty,"96 and it received unanimous support in the U.S. House of Representatives and near-unanimous support in the U.S. Senate.97 Professor Levin also notes that Congress has enabled religious objectors to Social Security taxes—notably including the Amish, one of whom lost a Free Exercise Clause challenge to such taxes in 198298—to apply for exemptions for themselves and their employees,99 and that in 2011 the executive branch also accommodated Amish religious beliefs.100

Finally, even in the highly regimented military profession, the legislative and executive branches have accommodated the religious practices of servicemembers. When Congress passed the National Defense Authorization Acts for fiscal years 2013 and 2014, it provided for the "[e]nhancement and protection of rights of conscience."101 Not long after, the Department of Defense (DOD) issued Instruction 1300.17(4)(a), which provides servicemembers with heightened free exercise protections. The Instruction states that "[t]he DOD places a high value on the rights of members of the Military Services to observe the tenets of their respective religions." Further, "[r]equests for religious accommodation will be resolved in a timely manner and will be approved," so long as they do not "adversely affect mission accomplishment."102 This Instruction was applied in the case of Iknor Singh, an observant Sikh who sought relief from the Army's uniform standards.103 The district court concluded that the Army failed to show that denying Singh a religious accommodation to observe his Sikh faith "further[ed] the government's compelling interests" or was "the least restrictive means of furthering [the government's] interests," both of which are required under the Instruction.104 As in Elauf’s case, the court protected a member of a minority religion by interpreting an already protective provision and applying it.

Through Army Directive 2017-03, the Army guaranteed even stronger protections for religious practices, specifically the practices of observant Sikhs. It directed “Army uniform and grooming policy to provide wear and appearance standards for the most commonly requested religious accommodations.”105 Simratpal Singh did not have the benefit of the Directive when he pursued “a permanent religious accommodation that would allow him to wear uncut hair, a beard, and a turban, as required by his Sikh faith, while serving in the Army.”106 A district court denied his attempt to obtain that permanent religious accommodation in light of the military’s generally stringent appearance and grooming standards. “Years of advocacy”—and likely court losses like his—inspired the issuance of the Directive in 2017.107

The Department of Veterans Affairs (VA), recognizing the religious diversity of its servicemembers, has funeral guidelines to honor each fallen soldier’s religious convictions.108

VA values and respects Veterans and their families’ right to committal services held at VA National Cemeteries...
that honor their faith tradition. The wishes of a deceased Veteran’s family remain paramount in determining what, if any, religious expression will take place at a Veteran’s committal service. Families are free to have a committal service with or without religious references or the display of religious or other symbols.

Furthermore, the Honoring America’s Veterans and Caring for Camp Lejeune Families Act of 2012 permits the placement of commemorative monuments in memory of an individual’s or group’s “service in the Armed Forces” in Arlington National Cemetery, and it does not prohibit the inclusion of religious symbols on those monuments.109

American Humanist Association brought up this Act during oral argument in the Peace Cross case, prompting Justice Samuel Alito to quip that its religiously neutral approach to memorializing was “the way this sort of thing is being handled today in a pluralistic society in which ordinary people get along pretty well and—and are not at each other’s throats about religious divisions.”110 Justice Alito’s comments capture the sentiment that accommodation for minority religious beliefs in a pluralistic society is available outside the courts; indeed, it is best to seek such accommodation outside the courts.

Of course, some will seek to take advantage of the political branches to exclude others from full participation in our society. No government institution—including the judiciary—can perfectly protect against human rivalry and selfishness. But as a matter of structure, the political branches have greater capacity to protect the rights of religious minorities and to respond to bad policy. After all, “[o]nce a court issues a ruling, the doctrine of stare decisis immediately encamps around it to stifle any later change or repudiation. That is not at all the situation with legislation, which can come and go as political power migrates from one set of interest groups to another.”111 The overall success of religious minorities in obtaining accommodations in legislation and executive action—and their mixed success and failure in the courts—shows that this is as true in practice as it is in theory.

IV. Conclusion

Religious minorities, like all Americans, want the law to protect their right to religious free exercise in the public square. An Establishment Clause doctrine that, in Thomas Jefferson’s words, reflects the clause’s meaning at the “time when the Constitution was adopted” and “recollect[s] the spirit manifested in the debates” benefits everyone by ensuring judicial objectivity and empowering the political branches to accommodate religious minorities.112 A historical approach for the courts and a reliance on the flexibility and responsiveness of the political branches is the best formula for a robust protection of religion—all religions—in the public square.

---


111 Kelsey, supra note 61, at 25.

112 Letter from Thomas Jefferson to Judge William Johnson, supra note 18, at 449.
The proliferation of distributed ledger technology, also known as blockchain, has the potential to disrupt or remake large sectors of the economy and is already doing so to some degree. An early application of blockchain began in 2008 with the introduction of Bitcoin, the well-known cryptocurrency. Following the emergence of Bitcoin, ambitious entrepreneurs and others began bringing to market their own digital currencies. And so the initial coin offering (ICO) was born.

An ICO is a form of financing in which an enterprise seeks to raise capital by selling a “coin” (sometimes called a “token”); the coin in turn gives the purchaser some future right in the business or other benefit or use. The interests of the coin holder are usually reflected in an electronic smart contract, and ownership of the coin is reflected on a digital ledger. The term “ICO” is a riff on IPO, or initial public offering.

As these offerings have become more common, a wide variety of terms have been deployed to describe the underlying asset being offered: coin, token, cryptocurrency, digital currency, digital asset, and crypto asset, to name a few. In some cases, the terms are used interchangeably, and in others, people differentiate among them purposefully to highlight subtle nuances in form or substance. Indeed, even the term ICO has waned in some circles, and people have turned to the phrase “security token” (or some variant thereof) to connote an offering that is subject to, and thus must comply with, the federal securities laws. Sometimes, the term “utility token” is used if a token’s value resides in its functionality, indicating that it therefore is not an investment subject to the federal securities laws and the jurisdiction of the U.S. Securities and Exchange Commission (SEC).

Whatever they are called, ICOs (a phrase we will use flexibly in this article) have spurred debate over the potential application of the federal securities laws. The SEC has asserted oversight over this burgeoning market when a security is offered. This raises the central question: When is a coin or token a security? To determine whether and how to regulate this twenty-first century innovation, the SEC has sought guidance from the past—a 1946 Supreme Court case about orange groves.

This article addresses several key regulatory developments at the SEC that are influencing the shape of the crypto market. In particular, through recent announcements and enforcement actions, the agency has indicated when it believes a coin or token is a security subject to its jurisdiction. At the end of this article, we reference some of the relevant regulatory and enforcement efforts of other federal regulators and note that state securities regulators, which share anti-fraud and other authority with the SEC, have also been active in policing ICO activity.

I. The Investment Contract

It starts with the definition of “security.” If a security is involved, the federal securities laws are triggered, in toto. If there is no security, then the SEC lacks jurisdiction over the instrument.
A fundamental tenet of federal securities regulation is found in Section 5 of the Securities Act of 1933 (Securities Act). This provision requires every offer or sale of securities to be registered with the SEC or exempt from such registration under one or more statutory exemptions. An “offer” is defined broadly under the Securities Act as “every attempt or offer to dispose of, or solicitation of an offer to buy, a security or interest in a security, for value.” There is usually no exemption available to an issuer when securities are distributed on a wide scale to large numbers of individuals, who are referred to as “retail investors.”

The definition of “security” contained in Section 2(a)(1) of the Securities Act and Section 3(a)(10) of its companion statute, the Securities Exchange Act of 1934 (Exchange Act), includes—in addition to familiar financial instruments like stocks and bonds—“investment contracts.” In the famous 1946 case, SEC v. Howey Co., the Supreme Court articulated the test for determining when an arrangement is an investment contract and, therefore, a security subject to the federal securities laws. The Howey test has been in use ever since.

The Howey Company owned land in Florida where it cultivated orange groves. To fund new development, Howey sought outside financing and turned to out-of-state tourists who visited a hotel that adjoined one of its properties. Each prospective purchaser was offered both a land sales contract and a service contract; under the service contract, a Howey affiliate would manage the land on the purchaser’s behalf. Since the offerees were primarily non-residents with no wherewithal to care for orange groves, most of those who purchased an interest in the groves also accepted the service contract arrangement.

Upon payment of the purchase price, the land was conveyed to the purchaser, but individual tracts were not separately fenced and were identified by land marks intelligible only through a plat book record. The service contract granted Howey a leasehold interest and “full and complete” possession of the land. For a specified fee plus the cost of labor and materials, Howey had full discretion and authority over the cultivation of the groves and the harvest and marketing of the crops. Without Howey’s consent, purchasers had no right of entry to market the crops. Instead, Howey allocated a share of net profits to each purchaser after the harvest.

The Supreme Court was asked to determine whether the land sales contract, the warranty deed, and the service contract together constituted an investment contract under Section 2(a)(1) of the Securities Act. The lower courts found that no investment contract existed, and instead treated the contracts and deeds as separate transactions involving a sale of real estate and an agreement by the seller to manage the property for the buyer.

The Supreme Court began its analysis by considering the historical background against which the federal securities laws were adopted, as well as Congress’ intent in enacting the Securities Act and including the concept of an investment contract. The Court reasoned that the term investment contract “embodies a flexible rather than a static principle, one that is capable of adaptation to meet the countless and variable schemes devised by those who seek the use of the money of others on the promise of profits.” In this view, a flexible understanding of what an investment contract is helps ensure that the reach of the federal securities laws is not unduly circumscribed.

The transactions in this case clearly involve investment contracts as so defined. The respondent companies are offering something more than fee simple interests in land, something different from a farm or orchard coupled with management services. They are offering an opportunity to contribute money and to share in the profits of a large citrus fruit enterprise managed and partly owned by respondents. They are offering this opportunity to persons who reside in distant localities and who lack the equipment and experience requisite to the cultivation, harvesting and marketing of the citrus products. Such persons have no desire to occupy the land or to develop it themselves; they are attracted solely by the prospects of a return on their investment. . . . A common enterprise managed by respondents or third parties with adequate personnel and equipment is therefore essential if the investors are to achieve their paramount aim of a return on their investments. Their respective shares in this enterprise are evidenced by land sales contracts and warranty deeds, which serve as a convenient method of determining the investors’ allocable shares of the profits. The resulting transfer of rights in land is purely incidental.

In short, the purchasers depended on Howey to run an orange business for them. The allocation of responsibility between Howey and the purchasers calls to mind the separation of ownership and control that characterizes the corporate form, where there exists a centralized management team and passive shareholders. Furthermore, as the Court explains, the purchasers did not buy interests in the land so that they could eat the produce that Howey grew. The purchasers’ motivation was investment, not consumption.

From this, we get the Howey test for determining whether an investment contract—and thus a security—exists. Under Howey, an investment contract exists if four factors are present:

(i) an investment of money by a person;
(ii) in a common enterprise;
(iii) where the person is led to expect profits;

1 Section 5 of the Securities Act generally requires an offeror of securities to register that offering with the SEC by means of a written filing (known as a “registration statement”) and deliver each offeror a prospectus containing various required disclosures. This process is time-consuming and can be expensive, typically requiring the assistance of experienced securities counsel, public accountants, and other professional advisors.

2 328 U.S. 293 (1946).

3 Id. at 299.

4 Id. at 299-300.

5 See generally James D. Cox & Thomas Lee Hazen, 1 Treatise on the Law of Corporations §2.7 (3d ed. 2010).
For more than seventy years, securities lawyers and the SEC, along with the courts, have applied Howey to analyze a wide array of financial arrangements—such as limited partnership interests,7 condominiums,8 sale-leasebacks of payphones,9 and life settlements10—to determine whether they are subject to regulation by the SEC.

II. The SEC Issues the DAO Report

As cryptocurrencies have appeared on the scene and grown in use, the SEC has unsurprisingly tackled the question of whether ICOs should be considered offerings of securities. The SEC has looked to Howey, which the Supreme Court crafted to be malleable to new facts and circumstances, for the answer.

On July 25, 2017, the SEC released a Report of Investigation (the “DAO Report”)11 under Section 21(a) of the Exchange Act involving an issuer of tokens known as The DAO. Section 21(a) grants the SEC broad discretion to investigate potential violations of the federal securities laws and “publish information concerning any such violations.” From time to time, the agency uses this power over the business without a traditional corporate hierarchy. The DAO Report amounts to high-level guidance from the SEC.

The DAO is an example of a “decentralized autonomous organization” that exists via smart contracts executed on a blockchain, described by the SEC as a “‘virtual’ organization embodied in computer code.” The German company that created The DAO, Slock.it, automated its corporate governance structures and purported to give holders of DAO Tokens decision-making power over the business without a traditional corporate hierarchy.

In 2016, The DAO completed an ICO of DAO Tokens valued at approximately $150 million. DAO Tokens, which granted certain voting and ownership rights, were offered for sale to the general public in exchange for the cryptocurrency Ether. The DAO intended to use the Ether generated in the ICO to fund projects that would provide DAO Token holders a return on their investment. DAO Token holders had the right to vote on certain corporate governance matters of The DAO, including which projects to fund and when to make distributions of profits to holders of the tokens. After the ICO, token holders could trade their DAO Tokens on online platforms supporting secondary market transactions.

The SEC applied Howey to these facts and found that DAO Tokens are investment contracts that qualify as securities and thus must be offered in accordance with Section 5 of the Securities Act or fall within an exemption to it. Notably, the SEC emphasized its view that case law calls for focusing on substance over form and that the economic realities of a transaction matter, not its name.

In undertaking its analysis, the SEC quickly dispensed with the first three prongs of the Howey test, determining that when DAO Token holders invested Ether in The DAO, they were investing money in a common enterprise with the reasonable expectation of profits. As with most cases applying Howey over the decades, the “efforts of others” prong was central. The SEC ultimately concluded that the Howey test was met—and so an investment contract, and therefore a security, was present—because the efforts of Slock.it and the so-called “Curators” of proposals were “essential” to the enterprise, and DAO Token holders’ voting rights were limited.

To support its conclusion, the SEC observed that Slock.it created The DAO, maintained its coding and website, engaged in marketing, and chose individuals (the Curators) to screen investment opportunities so only the best projects were presented to DAO Token holders for a vote. Additionally, the SEC determined that DAO Token holders did not have meaningful control over The DAO because their voting rights were limited to pre-selected projects and the rules of the voting structure incentivized voting in favor of proposals. Furthermore, because the DAO Token holders were so widely dispersed and anonymous, there was no way for them to join together to exercise meaningful control as a practical matter. In terms of governance, the SEC determined that token holders were more like passive corporate shareholders than partners who have real authority in a general partnership.

In addition, The DAO had emphasized...
to purchasers their ability to re-sell their tokens in the secondary market, which The DAO, according to the SEC, helped facilitate.

Treating an ICO as a securities offering has consequences for secondary trading as well as the initial sale. Section 3(a)(1) of the Exchange Act defines “exchange” broadly to include any organization or group that provides a marketplace for bringing together purchasers and sellers of securities or otherwise performs the generally understood functions of a stock exchange. The SEC used the DAO Report to give notice that each of the online platforms supporting the secondary market for DAO Tokens appeared to operate as an exchange that would have to register with the SEC under Section 5 of the Exchange Act, since no exemption from registration seemed to be available.

As for the regulatory philosophy that underpins the DAO Report, the SEC was blunt: “The automation of certain functions through this technology, ‘smart contracts,’ or computer code, does not remove conduct from the purview of the U.S. federal securities laws.”

III. MUNCHEE AND MORE

Since issuing the DAO Report, the SEC has initiated numerous enforcement actions against promoters of ICOs. A number of these actions involved old-fashioned Ponzi schemes or other frauds masquerading as token offerings and do not raise any novel securities law questions. But one notable early case involved a token issuer against which the SEC made no allegations of fraud. On December 11, 2017, the SEC issued a cease-and-desist order against Munchee Inc.15 after finding that the company’s ICO involved unregistered offers and sales of securities in violation of Section 5 of the Securities Act.

According to the SEC, Munchee sought to raise $15 million for its blockchain-based food review and social platform by selling digital tokens that could be used to buy and sell goods and services in the future through an iPhone app. At the time of the ICO, the Munchee “ecosystem” was not yet functional, but the company planned to develop it with the proceeds raised in the offering. Munchee and others promoting the ICO represented to individuals that the tokens could be expected to increase in value as the company implemented improvements to the app, and they said that the company would work to support a secondary market for the tokens. Indeed, according to the settlement order, Munchee and its agent promoted the ICO to people interested in investing in digital assets, which “primed” investors’ profit expectations. Drawing on the DAO Report, the SEC concluded that the tokens were securities in the form of investment contracts under Howey.

After being contacted by the SEC, Munchee halted its ICO and refunded investors’ money before any tokens were delivered. Due to Munchee’s cooperation and its quick action to end the ICO and return funds, the SEC chose not to impose a penalty. Although the SEC sometimes brings standalone Section 5 cases where there is no allegation of fraud, such cases are infrequent. By selecting Munchee for enforcement, the SEC telegraphed that the SEC’s efforts to police the ICO market would not be limited to cases involving material misstatements or omissions of information. This is noteworthy in part because, unlike fraud, violation of Section 5 is a strict liability offense.

In a pair of cases brought on November 16, 2018, the SEC settled charges with two ICO issuers who conducted unregistered securities offerings.16 Both issuers sold tokens to investors to raise funds. The SEC found that purchasers of each issuer’s tokens would have had a reasonable expectation of obtaining a future profit based on each company’s respective efforts, including building out an “ecosystem” and adding new functionality using the proceeds from the sale. One of the companies also committed to support the value of its tokens by controlling the token supply. The SEC explained that each issuer made efforts to facilitate secondary trading and that their promotional communications indicated the profit potential. Consistent with Howey and the DAO Report, the SEC concluded that both companies had offered securities without registering them with the SEC. Unlike with Munchee, the SEC assessed $250,000 penalties against each company and required them to compensate investors, register the offerings, and begin filing periodic reports with the SEC under the Exchange Act.17

IV. SENIOR SEC STAFF WEIGH IN

William Hinman, Director of the SEC’s Division of Corporation Finance, delivered a speech on June 14, 2018, providing further insight into how the SEC analyzes ICOs under the Howey test.18 He began his remarks by reiterating that Bitcoin is not a security. In a notable move, Hinman also indicated that the SEC staff does not view Ether as a security either in its “present state,” saying nothing about what Ether’s status might have been under Howey in its earlier state. Hinman emphasized that the decentralized nature of the networks underlying both Bitcoin and Ether would mean that applying the disclosure requirements


17 Some ICO promoters facing an enforcement action from the SEC have declined to settle and instead have opted to litigate, asserting that their particular tokens do not satisfy the Howey test. In one recent case, a federal district court initially declined to grant the SEC’s request for a preliminary injunction against the issuer, Blockvest. The court ruled that, given the stage of the litigation and that there were disputed issues of material fact, the court could not determine that there was a security. See SEC v. Blockvest, Case No.: 18CV2287-GPB(BLM) (S.D. Cal. Nov. 27, 2018), available at https://www.scribd.com/document/394382912/Blockvest-Ruling#from_embed (order denying the SEC’s motion for a preliminary injunction against the issuer, Blockvest). On reconsideration, and after the introduction of new evidence that the court found supported finding a security under Howey, the judge reversed his earlier decision and issued the preliminary injunction. See SEC v. Blockvest, Case No.: 18CV2287-GPB(BLM) (S.D. Cal. Feb. 14, 2019), available at https://www.sec.gov/litigation/litreleases/2019/order24400.pdf.

of the federal securities laws would serve little purpose. Hinman posited, “when the efforts of the third party are no longer a key factor for determining the enterprise’s success, material information asymmetries recede.” In those circumstances, such as when a network becomes “truly decentralized,” according to Hinman, “the ability to identify an issuer or promoter to make the requisite disclosures becomes difficult, and less meaningful.”

Director Hinman’s comments reflect the fact that the Securities Act is designed to get material information that a promoter knows into the hands of investors so that investors can make informed decisions.

Hinman also noted that “the analysis of whether something is a security is not static and does not strictly inhere to the instrument,” which seems to suggest that it is possible for a coin or token that is a security to cease being one. Does this imply that the Howey test should be administered periodically to see how the facts and circumstances prevailing at different times fare under the investment contract analysis? What exactly it will take, in the SEC’s view, for an instrument’s status to morph from security to non-security, as well as the precise regulatory and practical implications of any such change, is uncertain.

Hinman concluded his speech with two sets of questions that go to the characterization of digital assets. The first deals with whether a third party “drives” (to use Hinman’s word) any expectation of profits that purchasers may have. The second set of questions is about whether a digital asset is consumable. As to factors to consider in assessing the efforts of others, Hinman asks:

- Is there a person or group that has sponsored or promoted the creation and sale of the digital asset, the efforts of whom play a significant role in the development and maintenance of the asset and its potential increase in value?
- Has this person or group retained a stake or other interest in the digital asset such that it would be motivated to expend efforts to cause an increase in value in the digital asset? Would purchasers reasonably believe such efforts will be undertaken and may result in a return on their investment in the digital asset?
- Has the promoter raised an amount of funds in excess of what may be needed to establish a functional network, and, if so, has it indicated how those funds may be used to support the value of the tokens or to increase the value of the enterprise? Does the promoter continue to expend funds from proceeds or operations to enhance the functionality and/or value of the system within which the tokens operate?
- Are purchasers “investing,” that is seeking a return? In that regard, is the instrument marketed and sold to the general public instead of to potential users of the network for a price that reasonably correlates with the market value of the good or service in the network?
- Does application of the Securities Act protections make sense? Is there a person or entity others are relying on that plays a key role in the profit-making of the enterprise such that disclosure of their activities and plans would be important to investors? Do informational asymmetries exist between the promoters and potential purchasers/investors in the digital asset?
- Do persons or entities other than the promoter exercise governance rights or meaningful influence?

As to factors that speak to consumption versus investment:

- Is token creation commensurate with meeting the needs of users or, rather, with feeding speculation?
- Are independent actors setting the price or is the promoter supporting the secondary market for the asset or otherwise influencing trading?
- Is it clear that the primary motivation for purchasing the digital asset is for personal use or consumption, as compared to investment? Have purchasers made representations as to their consumptive, as opposed to their investment, intent? Are the tokens available in increments that correlate with a consumptive versus investment intent?
- Are the tokens distributed in ways to meet users’ needs? For example, can the tokens be held or transferred only in amounts that correspond to a purchaser’s expected use? Are there built-in incentives that compel using the tokens promptly on the network, such as having the tokens degrade in value over time, or can the tokens be held for extended periods for investment?
- Is the asset marketed and distributed to potential users or the general public?
- Are the assets dispersed across a diverse user base or concentrated in the hands of a few that can exert influence over the application?
- Is the application fully functioning or in early stages of development?

All of this is summed up in one overarching question that Hinman poses to frame his speech: “But what about cases where there is no longer any central enterprise being invested in or where the digital asset is sold only to be used to purchase a good or service available through the network on which it was created?”

At a November 2018 conference, Hinman remarked that, in 2019, the SEC intends to provide some clearer answers to these and other relevant questions in the form of further guidance. One possibility for the guidance would be to consolidate SEC views into a sort of how-to manual for people to use in assessing the applicability of the federal securities laws.

V. On “Utility Tokens”

For decades, it has been widely acknowledged that the SEC regulates investment, not consumption. In 1975, in United Housing Foundation v. Forman, the Supreme Court held that an instrument, even though it was called “stock,” was not a security in the form of stock or an investment contract under the federal

---

VI. The “Airdrop”

If the instrument issued in an ICO is a security that requires SEC registration in the absence of an exemption, can an issuer avoid the registration requirements by simply giving coins or tokens away? Under certain circumstances, a bona fide gift of securities is deemed not to involve the offer or sale of those securities, and under this “no sale” theory the registration requirements of the Securities Act do not apply. In the ICO context, an “airdrop” generally refers to the widespread distribution of digital tokens to community members either for free or in exchange for performing minor tasks. The SEC addressed an airdrop used to distribute digital tokens to investors in an August 2018 enforcement action against an ICO issuer, Tomahawk Exploration LLC, and its promoter.26

According to the SEC, Tomahawk sought to raise $5 million through an ICO, said to fund oil drilling in California. When it failed to raise any money, the company instead made an airdrop of tokens to third parties by means of what it called a “bounty program” in exchange for online promotional and marketing services that targeted potential investors and directed them to the company’s offering materials. Following its DAO Report, the SEC concluded that the Tomahawk tokens were securities. The SEC also alleged a series of materially false and misleading statements in Tomahawk’s marketing documents.

The SEC then analyzed the company’s bounty program. The SEC determined that the company’s issuance of tokens under the bounty program constituted an offer and sale of securities because Tomahawk provided tokens to investors in exchange for services designed to advance the company’s economic interests and foster a trading market for its securities. The SEC reasoned that the lack of monetary consideration for “free” shares did not mean there was not an offer or sale under the federal securities laws. Rather, according to the SEC, a “gift” of a security is a “sale” for securities law purposes when the company receives some real benefit, even if it does not involve the exchange of money.

20 421 U.S. 837 (1975). Justices Brennan, Douglas, and White dissented, arguing that the shares were both stock and investment contracts under the federal securities laws.

21 See id. at 852-53 (“[W]hen a purchaser is motivated by a desire to use or consume the item purchased—to occupy the land or develop it themselves,” as the Howey court put it . . . — the securities laws do not apply.”); id. at 858 (“What distinguishes a security transaction . . . is an investment where one parts with his money in the hope of receiving profits from the efforts of others, and not where he purchases a commodity for personal consumption or living quarters for personal use.”). In addition, as the Court explained, the shares of stock at issue “cannot be transferred to a non-tenant; nor can they be pledged or encumbered; and they descend, along with the apartment, only to a surviving spouse. No voting rights attach to the shares as such: participation in the affairs of the cooperative appertains to the apartment, with the residents of each apartment being entitled to one vote irrespective of the number of shares owned.” Id. at 842. Stock also could not be re-sold at a profit. Id. at 843, 22

22 22  See id. at 851-52 (finding that the “efforts of others” prong of Howey is not met. Some cases decided before the proliferation of digital assets have found that there is no investment contract where an increase in the price of an instrument is the result of market forces or some other extrinsic factor, and not the promoter’s managerial

or entrepreneurial efforts.24 Or, as cases applying Howey to partnership and limited liability company interests have held, no investment contract exists where the holders of the interests exercise sufficient (even if not total) control over the enterprise or otherwise meaningfully participate in the business operations.25

24 See, e.g., McCown v. Heidler, 527 F.2d 204, 208 (10th Cir. 1976) (explaining that “land, as such, is not a security and that a land purchase contract, simply because the purchaser expects or hopes that the value of the land purchased will increase, does not fall automatically within the confines of the Securities Acts”). See also Noa v. Key Futures, Inc., 638 F.2d 77, 79 (1980) (9th Cir. 1980) (finding that a Contract of Purchase and a Confirmation and Certificate of Ownership concerning the sale of silver did not create an investment contract, explaining that “[o]nce the purchase of silver bars was made, the profits to the investor depended upon the fluctuations of the silver market, not the managerial efforts of Key Futures”).

25 See generally LOSS ET AL., supra note 6 at 1106-27.

The SEC found that Tomahawk received value in exchange for the bounty distributions in the form of online marketing, including the promotion of the ICO on blogs and other online forums. The company also received value in the creation of a public trading market for its securities. Accordingly, the SEC determined that the company issued tokens as part of the bounty program to generate interest in the ICO, which in turn benefited the company. Thus, the SEC concluded that a sale had occurred without registration in violation of Section 5 of the Securities Act.

This case is reminiscent of the SEC’s enforcement actions against several internet companies that distributed “free stock” during the height of the dot-com era twenty years ago. In what have become known to securities lawyers as the free stock cases, investors were typically required to sign up on issuers’ websites and disclose personal information in order to obtain “free” shares. Free stock recipients were also offered extra shares for soliciting additional investors or for linking their own websites to those of an issuer or purchasing services offered through an issuer. Due to these activities, the SEC similarly took the position that the issuers received value (and did not make a gift) by creating a public market for their shares, increasing their business prospects, creating publicity, increasing traffic to their websites, and generating possible interest in future securities offerings.

Call it an airdrop or call it free stock, the SEC continues to focus on a transaction’s substance, not its label.

VII. Beyond the Regulation of Securities Offerings

If securities are involved, the entirety of federal securities regulation is in play, including the requirement that broker-dealers and investment companies register with the SEC.

On September 11, 2018, the SEC announced its first case charging unregistered broker-dealers for selling digital tokens. According to the SEC’s order, the defendants operated a self-described “ICO Superstore” that solicited investors, took thousands of customer orders for digital tokens, processed investor funds, and handled more than 200 different digital tokens in connection with both ICOs and the defendants’ own secondary market activities.27 The defendants also promoted the sale of approximately forty digital tokens in exchange for marketing fees paid by digital token issuers. Because the digital tokens issued in the ICOs and traded by defendants included securities, the SEC concluded that the defendants’ activities required broker-dealer registration with the SEC.

The same day, the SEC also announced charges against a digital asset fund manager who failed to register the fund it advised with the SEC and misrepresented the manager’s status as a regulated entity.28 The SEC’s order cites the Investment Company Act of 1940 (Investment Company Act), which defines “investment company” as any issuer who:

- is engaged or proposes to engage in the business of investing, reinvesting, owning, holding or trading in securities, and owns or proposes to acquire investment securities having a value exceeding 40 percent of the value of such issuer’s total assets (exclusive of government securities and cash items) on an unconsolidated basis.

The SEC concluded that the fund engaged in the business of investing, holding, and trading digital assets that were securities and therefore had to register as an investment company under the Investment Company Act. The SEC also concluded that the fund manager, as an investment adviser, made material misstatements and omissions in violation of the Investment Advisers Act of 1940, as well as the Securities Act.

The Financial Industry Regulatory Authority (FINRA), which oversees broker-dealers, announced its first disciplinary action involving cryptocurrencies against a broker-dealer registered representative, again on September 11.29 According to FINRA’s complaint, the respondent attempted to attract investors into a penny stock company he controlled by offering interests in what he advertised as the “the first minable coin backed by marketable securities.” FINRA alleged that, as the company’s business struggled, the respondent acquired the rights to a cryptocurrency named HempCoin and attempted to repackage HempCoin as a security backed by the publicly traded penny stock. The respondent also marketed HempCoin as “the world’s first currency to represent equity ownership” in a publicly traded company. FINRA said that investors mined more than 81 million HempCoin through late 2017 and traded the security on two cryptocurrency exchanges. Based on this, FINRA alleged that the respondent engaged in the unlawful distribution of HempCoin as an unregistered security, made several misrepresentations, and never disclosed these transactions to his broker-dealer employer. Thus, FINRA asserted that the individual violated not only the federal securities laws, but also several FINRA regulations, including one requiring that registered representatives must “observe high standards of commercial honor and just and equitable principles of trade.”

VIII. Unregistered Token Exchanges

On November 8, 2018, the SEC announced settled charges against an unlicensed digital token platform called EtherDelta.30 The case is the SEC’s first enforcement action based on findings that such a platform operated as an unregistered national securities exchange. According to the SEC’s order, EtherDelta provided online secondary market trading of ERC20 tokens, a type of blockchain-based token commonly issued in ICOs. The SEC found that almost all of the orders placed through EtherDelta were traded after the SEC issued the DAO Report, which had mentioned that the federal securities laws provide a functional

---

test that could include a digital asset trading system within the definition of an “exchange” subject to the SEC’s jurisdiction.31

The case is particularly significant because the platform operated on a decentralized basis through programming in its smart contract that runs on a blockchain. The SEC found that EtherDelta’s smart contract was coded to validate the order messages, confirm the terms and conditions of orders, execute paired orders, and direct a distributed ledger to be updated to reflect a trade. The SEC also found that the individual behind EtherDelta caused the platform’s Exchange Act violation because he wrote and deployed the smart contract and controlled EtherDelta’s operations.

IX. Joint Statement by SEC Staff

Perhaps to highlight the growing emphasis on ICO enforcement, the SEC’s three principal rule-making divisions issued a joint statement (the Staff Statement) on November 16, 2018, summarizing many of the enforcement cases discussed above.32 The Staff Statement essentially reiterates the agency’s definition of an “exchange” subject to the SEC’s jurisdiction.31

The primary exemption from registration and regulation as an exchange is for so-called alternative trading systems (ATSs). Rule 3a1-1(a)(2) under the Exchange Act exempts from the definition of “exchange” any organization, association, or group of persons that complies with Regulation ATS. Regulation ATS, in turn, requires an ATS to, among other things, register as a broker-dealer, file a Form ATS with the SEC, and establish written safeguards and procedures to protect subscribers’ confidential trading information. An ATS that complies with Regulation ATS and otherwise complies with other applicable SEC regulations need not register as a national securities exchange.

31 Under the test in Exchange Act Rule 3b-16, a platform that does the following is treated as an exchange: (1) brings together the orders for securities of multiple buyers and sellers; and (2) uses established, non-discretionary methods (whether by providing a trading facility or by setting rules) under which such orders interact with each other, and the buyers and sellers entering such orders agree to the terms of the trade. The primary exemption from registration and regulation as an exchange is for so-called alternative trading systems (ATSs). Rule 3a1-1(a)(2) under the Exchange Act exempts from the definition of “exchange” any organization, association, or group of persons that complies with Regulation ATS. Regulation ATS, in turn, requires an ATS to, among other things, register as a broker-dealer, file a Form ATS with the SEC, and establish written safeguards and procedures to protect subscribers’ confidential trading information. An ATS that complies with Regulation ATS and otherwise complies with other applicable SEC regulations need not register as a national securities exchange.


33 The Staff Statement’s discussion of exchange registration goes beyond summarizing prior enforcement actions and relevant rules and regulations. It states that “an entity that provides an algorithm, run on a computer program or on a smart contract using blockchain technology, as a means to bring together or execute orders could be providing a trading facility” and that “an entity that sets execution priorities, indirectly, to provide various functions of a trading system that together meet the definition of an exchange, the entity arranging the collective efforts could be considered to have established an exchange.” How the staff considers treating a longstanding concept like an exchange in the context of a new technology like blockchain is instructive.

X. Bitcoin Exchange-Traded Funds

An exchange-traded fund, or ETF, is a generic term people use for a security that tracks a stock index or other basket of assets such as bonds or commodities.34 ETF shares trade on an exchange, but are otherwise very similar to mutual funds. Several entrepreneurs have recently conceived of Bitcoin-based ETFs (or other similar exchange-traded products), which under the federal securities laws cannot begin trading until they receive SEC approval.

In a lengthy order issued on July 26, 2018, by a 3-1 vote the SEC denied an application by the Bats BZX Exchange, Inc. (BZX) seeking to list and trade shares of the Winklevoss Bitcoin Trust.35 The Winklevoss brothers had been trying for two years to launch what would have been the first Bitcoin-based ETF in the U.S.36 As a threshold matter, BZX originally asserted that, for many investors, shares in the trust would represent a cost-effective and convenient means of gaining investment exposure to Bitcoin similar to a direct investment in Bitcoin.37 In support of its application, and to assuage potential SEC concerns around manipulation of the market for Bitcoin, BZX also argued that: (i) the “geographically diverse and continuous nature of bitcoin trading makes it difficult and prohibitively costly to manipulate the price of bitcoin,” and that, therefore, the Bitcoin market “generally is less susceptible to manipulation than the equity, fixed income, and commodity futures markets,” and (ii) “novel systems intrinsic to this new market provide unique additional protections that are unavailable in traditional commodity markets.”

In denying the application, the SEC cited various concerns about the lack of oversight in the underlying Bitcoin market and ruled that BZX did not demonstrate that Bitcoin and Bitcoin markets are adequately resistant to manipulation or that alternative means of detecting and deterring fraud and manipulation are sufficient in the absence of a surveillance-sharing agreement with a significant, regulated market related to Bitcoin. The SEC also stated that a substantial majority of Bitcoin trading occurs on unregulated venues overseas that are relatively new and that generally appear to trade only digital assets.


36 Although colloquially referred to as an “ETF” in much of the financial press, the SEC order technically classifies it as a “commodity-trust exchange-traded product.”

Furthermore, in the SEC’s view, regulated Bitcoin-related markets are still in the early stages of development, and the record did not support finding that Bitcoin derivatives markets have attained significant size. The SEC, therefore, concluded that BZX did not demonstrate that the structure of the spot market for Bitcoin is uniquely resistant to manipulation and likewise determined that current trading venues for Bitcoin are not resistant to market manipulation. More to the point, according to the SEC, BZX did not demonstrate, given the current absence of a surveillance-sharing agreement with a regulated Bitcoin market of significant size, that BZX’s proposed alternative surveillance procedures—including BZX’s claim that it could obtain information regarding trading in the trust shares and in the underlying Bitcoin or any Bitcoin derivative when needed—would satisfy the requirement that an exchange’s rules be designed to prevent fraud and manipulation.\(^{38}\)

Importantly, the SEC did not categorically rule out a Bitcoin ETF in the future. It left open the possibility that the Bitcoin market could grow and develop in ways that ameliorate the SEC’s concern about fraud and manipulation, or that other surveillance techniques could adequately mitigate the risk to investors.

In a vigorous dissent, Commissioner Hester Peirce argued that the BZX application satisfied the statutory standard and that the SEC should permit BZX to list and trade the Winklevoss product.\(^{39}\) She expressed deep concern that the denial of the application “undermines investor protection by precluding greater institutionalization of the bitcoin market.” Commissioner Peirce argued that more “institutional participation would ameliorate many of the Commission’s concerns with the bitcoin market that underlie its disapproval order.” More generally, she asserted that the majority’s “interpretation and application of the statutory standard sends a strong signal that innovation is unwelcome in our markets, a signal that may have effects far beyond the fate of bitcoin” ETFs.

On August 22, 2018, the SEC staff, acting under delegated authority from the Commission, denied applications for nine more Bitcoin ETFs. The orders denying applications by Cboe BZX\(^{40}\) and NYSE Arca\(^{41}\) are similar to each other and cite many of the same reasons for denial as those cited in the Winklevoss application. As with the Winklevoss disapproval order, the SEC staff emphasized that “its disapproval does not rest on an evaluation of whether bitcoin, or blockchain technology more generally, has utility or value as an innovation or an investment.” Instead, the SEC reasoned that the exchanges failed to meet their burdens under SEC regulations to demonstrate their ability to prevent fraudulent and manipulative acts and practices in respect of the planned ETFs. The SEC staff elaborated, finding that the exchanges did not demonstrate that Bitcoin futures markets are “markets of significant size.” The SEC staff explained that this is critical because the exchanges did not establish that other means to prevent fraudulent and manipulative acts and practices will be sufficient to prevent fraud; surveillance-sharing with a regulated market of significant size related to Bitcoin is therefore necessary, according to the staff, to satisfy the statutory requirement that the exchanges’ rules be designed to address such misconduct. By August 24, 2018, the SEC announced that the commissioners would review the staff’s findings, and the denial of the nine ETFs was stayed.

Whether the SEC will approve a Bitcoin ETF in 2019 is one of the most anticipated developments in the cryptocurrency space. Entrepreneurs continue to file new applications with the SEC as they wait to see what will happen.\(^{42}\)

XI. WHAT’S NEXT?

The SEC has repeatedly asserted jurisdiction, in various ways, over digital assets. Rather than swim against the Howey tide, the most recent iteration of a coin or token offering has involved selling “security tokens.” Like other securities, security tokens can have a range of attributes concerning voting rights, economic returns, and other features. In a security token offering, the issuer has recognized that the instrument is a security and attempts to comply with the wide variety of SEC regulations discussed above. There are token sales in the process of registration with the SEC, presumably because no private placement exemption is available for those offerings. In his November 2018 remarks, Director Hinman even acknowledged a backlog of filings with the SEC by parties seeking to conduct registered offerings, and he observed that the staff is processing them carefully due to the unique issues they raise.

There is still considerable regulatory uncertainty over how the particularities of federal securities regulation will apply to specific, concrete facts and circumstances. The details matter, and perhaps more than anything, market participants would like further clarity. Many would prefer that additional guidance

---

38 In a letter addressed to two trade associations representing the securities industry, the SEC staff has also described several criteria that, if satisfied, may persuade the staff to support the application for a cryptocurrency ETF. See U.S. Securities and Exchange Commission, Staff Letter: Engaging on Fund Innovation and Cryptocurrency-related Holdings (Jan. 18, 2018), available at https://www.sec.gov/divisions/investment/noaction/2018/cryptocurrency-011818.htm


come through channels other than enforcement actions. Speeches like Director Hinman’s and pronouncements like the Staff Statement give useful insight into the staff’s thinking. Indeed, the DAO Report itself provided a valuable indication of the SEC’s intentions.

More guidance, whether from the staff or the Commission itself, will likely come in 2019 and beyond as the agency continues delving into nuances, getting more and more granular over time as it considers actual offerings and other real-world blockchain use cases for trading and holding securities. The big-picture challenge for the SEC is to make sure that regulatory rigidity does not impede important technological innovation that stands to benefit entrepreneurs, investors, and our capital markets overall, while ensuring that investor protection is not jeopardized. The objectives of federal securities regulation need to be met, but specific regulatory requirements also need to make practical sense for digital assets and blockchain technology. This balance is achievable through regulatory fine-tuning that is informed by constructive collaboration between the SEC and market participants, and the SEC’s outreach to date is commendable.43

But the SEC is not the only regulator that matters to the future of cryptocurrency. Whether or not a digital asset meets the definition of a security under the federal securities laws, other regulators may regulate it under their regulatory regimes. For example, in 2014, the Commodity Futures Trading Commission classified Bitcoin and other digital currencies as “commodities” covered by the Commodity Exchange Act and subsequently has brought several enforcement actions against parties involved in the sale of digital currency.44 The Financial Crimes Enforcement Network has issued guidance stating that digital currency is considered currency and that exchanges will be considered exchanges under the Bank Secrecy Act.45 The Federal Trade Commission has established a blockchain working group and has brought enforcement actions against promoters of allegedly fraudulent chain referral schemes involving cryptocurrencies.46 A Congressional Blockchain Caucus has formed, and various members of Congress have introduced bills that would either expand or contract federal oversight of the space or simply usher in different regulation. And while this article has not addressed state securities regulation, federal law often does not preempt states’ authority in this area. Many states are now coordinating their enforcement efforts to pursue cases against fraudulent or unregistered ICOs.47

As this plays out, one concern is that entrepreneurs will conduct their offerings and other business offshore if U.S. regulation is overly restrictive and burdensome compared to the regulation in foreign jurisdictions. Making sure that the U.S. does not miss out on key blockchain developments and economic opportunity counsels in favor of ensuring that the U.S. regulatory environment, at both the federal and state levels, does not chill beneficial innovation.

For securities lawyers, this is a rare time. In the SEC’s 85-year history, no other development that has evolved so quickly, been the subject of so much varied regulatory attention, and held so much promise as digital assets and blockchain technology. And it all started with Howey’s orange groves.

43 In an effort to increase dialogue with the fintech community, the SEC has established a Strategic Hub for Innovation and Financial Technology (known as FinHub), with a website at https://www.sec.gov/finhub.


47 Since April 2018, for example, state and provincial securities regulators across the U.S. and Canada have been coordinating their ICO enforcement actions under “Operation Cryptosweep.” See generally North American Securities Administrators Association, Operation Cryptosweep, http://www.nasaa.org/regulatory-activity/enforcement-legal-activity/operation-cryptosweep.
The Americans with Disabilities Act (ADA)—enacted in 1990 and amended in 2008—was the first comprehensive federal civil rights law protecting individuals with disabilities. Title I of the ADA prohibits discrimination in employment and is enforced by the Equal Employment Opportunity Commission (EEOC). Title II “applies to state and local government entities, and protects qualified individuals with disabilities from discrimination on the basis of disability in services, programs, and activities provided by state and local government entities.”

Title III prohibits discrimination on the basis of disability by places of public accommodation (places that are privately owned, leased, or operated, and that affect commerce) that fall into one of twelve categories listed in the statute. Title III also “requires newly constructed or altered places of public accommodation— as well as commercial facilities (privately owned, nonresidential facilities such as factories, warehouses, or office buildings)—to comply with the ADA Standards.” Further, Title III covers “examinations and courses related to applications, licensing, certification, or credentialing for secondary or postsecondary education, professional, or trade purposes.” The United States Department of Justice (DOJ) enforces Title III (and portions of Title II) of the ADA.

The ADA defines “disability” as: “(A) a physical or mental impairment that substantially limits one or more major life activities of [an] individual; (B) a record of such an impairment; or (C) being regarded as having such an impairment.” The statute also sets out an expansive definition of “Major Life Activities” as well as rules of construction to interpret broadly the phrase “substantially limited.”

Central to the ADA is the requirement that entities subject to Title I provide reasonable accommodations to qualified disabled individuals unless doing so would cause an “undue hardship.”

Title III contains similar provisions that require public accommodations to make reasonable modifications to facilities,

---

1. See 42 U.S. Code § 12101, et seq. The focus of this article is on whether websites are covered under Title III of the ADA. It does not address state or local disability laws that may be applicable to website accessibility nor address possible defenses to website accessibility claims under Title III.

2. See 42 U.S. Code §§ 12111-12117. For the EEOC’s regulations implementing Title I of the ADA, see 29 C.F.R. Part 1630.


6. 42 U.S. Code § 12102(1).

7. 42 U.S. Code §§ 12102(2), (4).

8. 42 U.S. Code §§ 12111(8)-(10).
that are structural in nature may constitute discrimination under failure to remove architectural barriers or communication barriers that are structural in nature may constitute discrimination under Title III where removal is readily achievable.10

I. Are Websites Covered Under Title III?

In 1991, the DOJ enacted regulations to implement Titles II and III of the ADA, which were revised in September 2010.11 To qualify as a public accommodation under DOJ Title III regulations, an entity must fall within at least one of twelve categories:12

1. Places of lodging (e.g., inns, hotels, motels) (except for owner-occupied establishments renting fewer than six rooms);
2. Establishments serving food or drink (e.g., restaurants and bars);
3. Places of exhibition or entertainment (e.g., motion picture houses, theaters, concert halls, stadiums);
4. Places of public gathering (e.g., auditoriums, convention centers, lecture halls);
5. Sales or rental establishments (e.g., bakeries, grocery stores, hardware stores, shopping centers);
6. Service establishments (e.g., laundromats, dry cleaners, banks, barbershops, beauty shops, travel services, shoe repair services, funeral parlors, gas stations, offices of accountants or lawyers, pharmacies, insurance offices, professional offices of health care providers, hospitals);
7. Public transportation terminals, depots, or stations (not including facilities relating to air transportation);
8. Places of public display or collection (e.g., museums, libraries, galleries);
9. Places of recreation (e.g., parks, zoos, amusement parks);
10. Places of education (e.g., nursery schools, elementary, secondary, undergraduate, or postgraduate private schools);
11. Social service center establishments (e.g., day care centers, senior citizen centers, homeless shelters, food banks, adoption agencies); and
12. Places of exercise or recreation (e.g., gymnasiums, health spas, bowling alleys, golf courses).

The ADA was enacted prior to widespread use of the internet by individuals and businesses. Therefore, Title III and DOJ regulations do not specifically address the internet or provide guidelines for website compliance.

A. DOJ’s Past Position on Website Coverage Under Title III

In 1996, the Assistant Attorney General for Civil Rights, Deval Patrick, authored a letter in response to an inquiry from Sen. Tom Harkin. Patrick’s letter stated that covered entities under the ADA must make their websites accessible to provide effective communication.13 Thereafter, the DOJ actively pursued enforcement of website compliance with Title III through litigation which resulted in consent decrees, the filing of amicus briefs, and statements of interest.14

On July 26, 2010, the DOJ issued advance notice of proposed rulemaking to establish accessibility standards for website compliance.15 However, on December 26, 2017, the DOJ placed that rulemaking on the 2017 Inactive Actions list with no further information, although it stated it would “continue to assess whether specific technical standards are necessary and appropriate to assist covered entities with complying with the ADA.”16

In 2017, the DOJ appeared to change its view on whether websites are covered under the ADA in an amicus brief filed with the U.S. Supreme Court opposing certiorari in Magee v. Coca-Cola Refreshment USA, where the Fifth Circuit had held that a vending machine was not a place of public accommodation.17 In its amicus brief, the DOJ argued that “the court of appeals correctly held that the beverage vending machines at issue are not ‘place[s] of public accommodation’ under Title III of the ADA.”18 The DOJ further argued that “questions concerning Title III’s application to nonphysical establishments—including websites or digital services—may someday warrant this Court’s attention . . . .

14 See id. (listing and describing cases).
15 Id.
17 833 F.3d 530 (5th Cir. 2016), cert. denied, ___ U.S. ___ (Oct. 2, 2017) (No. 16-668).
case is not a suitable vehicle for addressing those emerging issues, however, since petitioned encounter respondent’s machines in person, not by telephone or over the Internet.”

**B. DOJ’s Current Position on Website Coverage Under Title III**

On September 4, 2018, Sen. Chuck Grassley wrote a letter to then-Attorney General Jeff Sessions, encouraging the DOJ to clarify whether the ADA applies to websites given the increase in lawsuits filed over alleged website inaccessibility.36 Earlier, on June 20, 2018, over 100 members of Congress had sent a letter to Sessions complaining about the lack of clarity for website compliance under the ADA in light of actual and threatened legal action by plaintiffs’ attorneys across the country.21

On September 25, 2018, Assistant Attorney General Stephen E. Boyd responded to the June 20 letter confirming DOJ’s earlier position that the ADA applies to the websites of public accommodations. He stated that the DOJ’s “interpretation is consistent with the ADA’s Title III requirement that the goods, services, privileges, or activities provided by places of public accommodation be equally accessible to people with disabilities.”22 Boyd also stated that, “absent the adoption of specific technical requirements for websites through rulemaking, public accommodations have flexibility in how to comply with the ADA’s general requirements of nondiscrimination and effective communication” and that “noncompliance with a voluntary technical standard for website accessibility does not necessarily indicate noncompliance with the ADA.”23

**II. Existing Website Accessibility Guidelines**

While the DOJ has not promulgated regulations setting forth guidelines for website accessibility under Title III, it has pointed out that:

"[T]he Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) has created recognized voluntary international guidelines for Web accessibility. These guidelines, set out in the Web Content Accessibility Guidelines (WCAG), detail how to make Web content accessible to individuals with disabilities... The WCAG 2.0 contains 12 guidelines addressing Web accessibility. Each guideline contains testable criteria for objectively determining if Web content satisfies the guideline. In order for a Web page to conform to the WCAG 2.0, the Web page must satisfy the criteria for all 12 guidelines under one of three conformance levels: A, AA, or AAA.24 W3C provides online resources and tools such as tutorials and support materials to assist organizations in making their websites accessible to disabled individuals."25

Federal agencies are subject to the Electronic and Information Technology Standards found in Section 508 of the Rehabilitation Act of 1973.26 These standards, known colloquially as section 508 standards, are published by the U.S. Access Board.27

**III. Illustrative Title III Website Accessibility Cases**

As referenced in the letters of Sen. Grassley and members of Congress to the Attorney General, the number of ADA Title III website accessibility lawsuits (as well as Title III lawsuits overall) has been increasing for the past two years. According to a recent legal blog post, plaintiffs filed at least 2,258 website accessibility lawsuits in 2018, a 177% increase from 814 such lawsuits in 2017.38 Only 262 website accessibility lawsuits were filed in 2015 and 2016 combined.29 New York (with 630) and Florida (with 342) led the country in website accessibility cases in the first half of 2018.30

**A. Circuit Split Identified by Recent Decision**

The state of the law on Title III website accessibility cases is evolving, and marked differences are developing among the circuit courts of appeal and even among district courts within the same circuit. The division among the courts on the issue of whether websites are a “public accommodation” was noted by the district court in *Gil v. Winn Dixie Stores, Inc.*:

“Courts are split on whether the ADA limits places of public accommodation to physical spaces. Courts in the First, Second, and Seventh Circuits have found that the ADA can apply to a website independent of any connection between

---

24 Nondiscrimination on the Basis of Disability ANPRM, supra note 13, at Section IV.A. See WCAG 2.0 Guidelines, http://www.w3.org/TR/WCAG20/.


26 29 U.S.C. 794d.


the website and a physical place. Courts in these circuits have typically looked at Congress's intent that individuals with disabilities fully enjoy the goods, services, privileges and advantages available indiscriminately to other members of the public, and at the legislative history of the ADA, which indicates that Congress intended the ADA to adapt to changes in technology.

On the other hand, courts in the Third, Sixth, and Ninth Circuits have concluded that places of public accommodation must be physical places, and that goods and services provided by a public accommodation must have a sufficient nexus to a physical place in order to be covered by the ADA. Courts in these circuits have concluded that a public accommodation must be a physical place because the 12 enumerated categories of public accommodations in the statute are all physical places.31

The district court in Gil noted that its own court of appeals, the Eleventh Circuit, had “not addressed whether websites are public accommodations for purposes of the ADA,” but had held that the ADA covers both tangible and intangible barriers to a disabled person’s ability to access public accommodations to goods and services.32

**B. Eleventh Circuit Provides Expansive Interpretation**

More recently, in *Haynes v. Dunkin’ Donuts LLC*, the Eleventh Circuit clarified its position on the ADA’s application to websites in reversing a Florida district court’s grant of a motion to dismiss a Title III website accessibility lawsuit.33 Plaintiff Dennis Haynes is blind and uses screen reading software (a program called JAWS) to navigate the internet.34 He claimed that Dunkin’ violated his rights under Title III because its website is incompatible with screen reader software, Haynes alleges that neither he, nor any blind person, can use those features.35

The Eleventh Circuit, in resolving the appeal in Haynes’ favor, noted that the ADA prohibits discrimination as to both tangible and “intangible barriers.”36 It further noted that Haynes had shown a plausible claim for relief under Title III with the following complaint allegations:

The inaccessibility of Dunkin’ Donuts’ website has similarly denied blind people the ability to enjoy the goods, services, privileges, and advantages of Dunkin’ Donuts’ shops. Among other things, he alleges that Dunkin’ Donuts’ website allows customers to locate physical Dunkin’ Donuts store locations and purchase gift cards online. Haynes also alleges that Dunkin’ Donuts’ website “provides access to” and “information about . . . the goods, services, facilities, privileges, advantages or accommodations of” Dunkin’ Donuts’ shops. Because the website isn’t compatible with screen reader software, Haynes alleges that neither he, nor any blind person, can use those features.38

The Eleventh Circuit ended its analysis by stating:

[It] appears that the website is a service that facilitates the use of Dunkin’ Donuts’ shops, which are places of public accommodation . . . and the ADA is clear that whatever goods and services Dunkin’ Donuts offers as a part of its place of public accommodation, it cannot discriminate against people on the basis of a disability, even if those goods and services are intangible.39

**C. Where Does the Sixth Circuit Stand?**

The Sixth Circuit has not directly weighed in on whether a website can be a public accommodation under the ADA, but several district courts have interpreted Sixth Circuit precedent in addressing website accessibility lawsuits. Recently, in *Brintley v. Aeroquip Credit Union*, a district court in Michigan denied Aeroquip Credit Union’s motion to dismiss the plaintiff’s website accessibility claim under Title III and Michigan law.40

Plaintiff Karla Brintley, who is permanently blind, alleged that she was unable to effectively access Aeroquip’s website with her screen reader, which “hindered her from effectively browsing for locations, amenities, and services and deterred her from visiting Defendant’s branches.”41 In its motion to dismiss, Aeroquip argued that Brintley lacked standing to assert a claim, since she was not eligible to join the credit union, and that she had failed to state a claim.42 The district court rejected Aeroquip’s standing argument, finding that “eligibility for membership in the credit union [wa]s not a prerequisite for standing” based on the Supreme Court’s decision in *PGA Tour, Inc. v. Martin*,43 which noted that Title III does not contain a “clients or customers limitation.”44 The district court further found that “the barriers Plaintiff encountered when she tried to access Defendant’s website constitute a concrete

---


33 2018 WL 3634720 (11th Cir. 2018).

34 Id. at *1.

35 Id.

36 Id.

37 Id. at *2.

38 Id. (citing 42 U.S.C. § 12182(a)).

39 Id. (citing 42 U.S.C. § 12182(a)).


41 Id. at 788.

42 Id.


44 Aeroquip Credit Union, 321 F. Supp. 3d at 790.
and particularized injury for purposes of establishing Article III standing.45

After dispensing with the standing issue, the district court considered whether Brintley had stated a claim for relief under the ADA. First, the court noted that a credit union is a place of public accommodation, but recognized the split of authority on whether Title III applies to websites connected to a place of public accommodation.46 Upon a review of Sixth Circuit Title III case law,47 the district court noted that the Sixth Circuit has held that Title III only applies to “physical places” of public accommodation, but it rejected Aeroquip’s argument that those holdings “precluded relief under Title III for all claims concerning websites.”48 The district court stated that the Sixth Circuit “expressed no opinion as to whether a plaintiff must physically enter a public accommodation to bring suit under Title III as opposed to merely accessing, by some other means, a service or good provided by a public accommodation.”49 It pointed out that other courts had characterized the Sixth Circuit’s approach to Title III as a “nexus theory,” where Title III is violated “if the discriminatory conduct [the inaccessible website] has a ‘nexus’ to the goods and services of a physical location.”50

Relying on the reasoning of the district court in Castillo v. Jo-Ann Stores, LLC,51 the court found that “the Complaint sufficiently alleges a nexus between Defendant’s website and its brick-and-mortar locations,” further noting that the “website provides goods and services including a store locator, descriptions of amenities, and information about the services Defendant offers.”52 The court found that the access barriers to Aeroquip’s website “deterred Plaintiff from visiting Defendant’s physical locations.”53 The court concluded by rejecting Aeroquip’s argument that Brintley’s request for injunctive relief violated due process because neither the DOJ regulations nor Title III provide website accessibility guidelines.54

By contrast, in separate recent cases filed by a different plaintiff in the Northern District of Ohio against different credit union defendants, the district court never reached the website coverage issue. Instead, it granted the defendants’ motions to dismiss on standing grounds since the plaintiff was not eligible for membership in the credit union and failed to allege intent to use their services.55

D. Plaintiffs Prevail at Trial and Obtain Favorable Summary Judgment Ruling

Recently, a plaintiff prevailed in the first ADA Title III website accessibility case to go to trial. Following a non-jury trial, the court in Gil v. Winn-Dixie Stores, Inc. ruled that the plaintiff had standing to seek injunctive relief and found that the defendant violated Title III of the ADA by failing to maintain a website that was accessible to visually impaired customers. The court ordered the defendant to make modifications to the website so it would be accessible to visually impaired customers. In its ruling, the court stated that it:

[N]eed not decide whether Winn–Dixie’s website is a public accommodation in and of itself, because the factual findings demonstrate that the website is heavily integrated with Winn–Dixie’s physical store locations and operates as a gateway to the physical store locations. Although Winn–Dixie argues that Gil has not been denied access to Winn–Dixie’s physical store locations as a result of the inaccessibility of the website, the ADA does not merely require physical access to a place of public accommodation. Rather, the ADA requires that disabled individuals be provided “full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations of any place of public accommodation...” 42 U.S.C. § 12182(a). The services offered on Winn–Dixie’s website, such as the online pharmacy management system, the ability to access digital coupons that link automatically to a customer’s rewards card, and the ability to find store locations, are undoubtedly services, privileges, advantages, and accommodations offered by Winn–Dixie’s physical store locations. These services, privileges, advantages, and accommodations are especially important for visually impaired individuals since it is difficult, if not impossible, for such individuals to use paper coupons found in newspapers or in the grocery stores, to locate the physical stores by other means, and to physically go to a pharmacy location in order to fill prescriptions.

The factual findings demonstrate that Winn–Dixie’s website is inaccessible to visually impaired individuals who must use screen reader software. Therefore, Winn–Dixie has violated the ADA because the inaccessibility of its website has denied Gil the full and equal enjoyment of standing where plaintiff was not eligible to be a member of the defendant credit union and had no plans to become a member).
the goods, services, facilities, privileges, advantages, or accommodations that Winn–Dixie offers to its sighted customers.\textsuperscript{56}

In Gomez v. General Nutrition Corporation, the Southern District of Florida granted the plaintiff’s motion for summary judgment as to liability. The court found that the defendant’s “website is a place of public accommodation within the meaning of the ADA” since it “facilitates the use of the physical stores by providing a store locator.”\textsuperscript{57} The court further found that the website permitted products to be “purchased remotely” as “a service of the physical stores,” provided information about “promotions and deals,” and “operates as a gateway to the physical stores.”\textsuperscript{58} In short, the court found that the “inaccessibility of the website amounts to a denial of that service to blind individuals” and violates the ADA.\textsuperscript{59}

IV. Applicability to Websites of Online-Only Businesses

Several district courts have found the websites of online-only businesses to be public accommodations under Title III. For example, in National Federation of the Blind v. Scribd Inc., a blind plaintiff alleged that a digital subscription service library violated Title III of the ADA because its website and mobile applications were inaccessible to the blind.\textsuperscript{60} The district court, after finding the ADA ambiguous on the issue of whether a website qualified as a public accommodation, determined that the ADA’s legislative history compelled a finding that a public accommodation is not limited to a physical space and denied the defendant’s motion to dismiss.\textsuperscript{61}

Later, in Access Now, Inc. v. Blue Apron, LLC, a district court in New Hampshire denied the defendant’s motion to dismiss, relying on First Circuit precedent “that ‘public accommodations’ are not limited to actual, physical places.”\textsuperscript{62} The district court found the plaintiff’s complaint sufficiently pled a violation of Title III:

Applying the reasoning of Carparts to this case, the court cannot conclude, at the Rule 12(b)(6) stage, that the plaintiff’s complaint falls short of pleading that Blue Apron’s website is a “public accommodation” under Title III of the ADA. Though true that websites are not specifically mentioned in the twelve enumerated categories of “public accommodations,” the plaintiffs “must show only that the web site falls within a general category listed under the ADA.” Here, as Access Now argues, Blue Apron may amount to an online “grocery store,” which is listed under Title III’s definition of “public accommodation,” 42 U.S.C. § 12181(7)(E), or at the very least may fall within the general “other sales” or “other service establishment” categories, id. § 12181(7)(E)–(F). This suffices at the 12(b) (6) stage to prevent dismissal.\textsuperscript{63}

However, the district court acknowledged that courts in some other circuits require that a public accommodation be a physical space or have a nexus with a physical space.\textsuperscript{64}

V. Federal Legislation

In an effort to curb the surge in disability access cases under the ADA, the U.S. House of Representatives passed H.R. 620 in the 115th Congress in 2017. The bill would prohibit civil actions based on the failure to remove an architectural barrier to access into an existing public accommodation unless prior notice of the barrier is given to the owner/operator and the owner/operator fails to provide written notice of steps to be taken to improve the barrier or fails to remove or make substantial progress removing the barrier following the written description.\textsuperscript{65} However, the bill has not advanced in the U.S. Senate.

VI. Conclusion

While case law on website accessibility under Title III is still developing, courts applying the “nexus theory” have found violations where a public accommodation’s inaccessible website is closely integrated with its physical store location. This was evidenced in the Gil and Gomez rulings out of the Southern District of Florida and the Eleventh Circuit’s recent decision in Haynes. Some courts, as evidenced by the rulings in Scribd Inc. and Blue Apron, LLC, have found that the websites of online-only businesses may qualify as public accommodations. It is unlikely that DOJ will issue regulations or provide other official guidance in the near term, given its most recent actions on the issue, so it is

\textsuperscript{56} 257 F. Supp. 3d 1340, 1348 (S.D. Fla. 2017).
\textsuperscript{57} 323 F. Supp. 3d 1368, 1376 (S.D. Fla. 2018).
\textsuperscript{58} Id.
\textsuperscript{59} Id.
\textsuperscript{60} 97 F. Supp. 3d 565 (D. Vt. 2015).
\textsuperscript{61} Id. at 573-77. The court emphasized:

The fact that the ADA does not include web-based services as a specific example of a public accommodation is irrelevant because such services did not exist when the ADA was passed and because Congress intended the ADA to adapt to changes in technology. Notably, Congress did not intend to limit the ADA to the specific examples listed and the catchall categories must be construed liberally to effectuate congressional intent.

Id. at 571 (citing Nat’l Ass’n of the Blind v. Netflix, Inc., 869 F. Supp. 2d at 200-01 (D. Mass. 2012)). However, it also cited a number of contrary court decisions rejecting the argument that websites qualify as public accommodations. Id. at 569-70.

\textsuperscript{62} 2017 WL 5186354, at *3 (D.N.H. 2017) (quoting Carparts Distrib. Ctr., Inc. v. Auto. Wholesaler’s Ass’n of New England, Inc., 37 F.3d 12, 19 (1st Cir. 1994)).

\textsuperscript{63} Blue Apron, LLC, 2017 WL 5186354, at *4.

\textsuperscript{64} Id. (citing Magee, 833 F.3d at 534 (concluding that vending machines are not places of public accommodation because the ADA definition of public accommodation only includes actual physical spaces open to the public); Earll, 599 Fed. Appx. at 696 (“We have previously interpreted the term ‘place of public accommodation’ to require ‘some connection between the good or service complained of and an actual physical place.’”) (quoting Weyer, 198 F.3d at 1114); Ford, 145 F.3d at 612–14 (rejecting the reasoning in Carparts and holding that “public accommodation” does not refer to non-physical access); Parker, 121 F.3d at 1013–14 (“The clear connotation of the words in § 12181(7) is that a public accommodation is a physical place.”)).

reasonable to expect a continued increase in website accessibility cases under Title III.
Is It Time to Revisit the Constitutionality of Unauthorized Practice of Law Rules?

By Michael E. Rosman

Every jurisdiction has a rule against the unauthorized practice of law. Rule 49 of the Rules of the District of Columbia Court of Appeals, for example, governs the unauthorized practice of law in the District of Columbia. Like virtually every other such rule in jurisdictions throughout the United States, it prohibits certain conduct by both non-lawyers and lawyers who are admitted to other bars, but not the D.C. Bar.

Two recent opinions—a concurrence joined by three judges of the Ohio Supreme Court and a United States Supreme Court opinion—have raised constitutional questions about regulations of the unauthorized practice of law. In this article, I briefly review the two cases and then identify aspects of unauthorized practice rules that might be subject to challenge in the future based on the analysis in these decisions.

I. In re Jones

The more obviously relevant opinion is In re Jones, which the Ohio Supreme Court decided on October 17, 2018.1 Alice Jones was licensed to practice law in Kentucky and applied for admission to the Ohio Bar in October 2015. The month after she applied for admission, she moved to Cincinnati and transferred to the Cincinnati office of her law firm.2 From that office, she practiced law exclusively in matters related to proceedings or potential proceedings in Kentucky.3 The Ohio Board of Commissioners on Character and Fitness concluded that Ms. Jones was engaging in the unauthorized practice of law in Ohio and thus violated Ohio’s Rule of Professional Conduct 5.5.4 It specifically rejected Jones’s claim that she was not engaging in the practice of law in Ohio because her presence there was temporary.5 It recommended that the Ohio Supreme Court disapprove of Jones’s application for admission to the bar based on her violation of this rule and her failure to provide clear evidence of her character and fitness.6

2 Id. at *2.
3 Id.
4 Id. Ohio Rule of Professional Conduct 5.5(b) provides that “[a] lawyer who is not admitted to practice in this jurisdiction shall not . . . except as authorized by these rules or other law, establish an office or other systematic and continuous presence in this jurisdiction for the purpose of practicing law.”
5 Jones, 2018 WL 5076017, at *2. Ohio Rule of Professional Conduct 5.5(c)(2) provides that an attorney admitted in another United States jurisdiction can provide legal services in Ohio on a temporary basis if the services are reasonably related to a proceeding or potential proceeding before a tribunal provided the lawyer was admitted before the tribunal. Other provisions of Rule 5.5(c) similarly provide that temporary legal services can be rendered in Ohio in negotiations, investigations, arbitration, mediation, and other non-litigation activity reasonably related to the lawyer’s practice in another jurisdiction.
6 Jones, 2018 WL 5076017, at *3.
The Ohio Supreme Court rejected the board’s recommendation and approved Jones's application for admission. It concluded that her Kentucky legal work from an Ohio office, even for the several years that her application was pending, was the provision of temporary legal services permitted by the rules.

Three justices concurred, even though they agreed with the board that Jones’s practice was not temporary because she had established an office. The concurring justices found, however, that the prohibition on Jones’s practicing Kentucky law from an Ohio office violated the Ohio and U.S. constitutions. They found that Ohio had no legitimate interest in regulating Jones’s representation of Kentucky clients in Kentucky tribunals simply because she maintained an office in Ohio. Ohio’s only legitimate interests in regulating the unauthorized practice of law—supervising the administration of justice in the state and protecting the Ohio public—are not served “when applied to a lawyer who is not practicing in Ohio courts or providing Ohio legal services, the Ohio rule on unauthorized practice violates the Ohio Constitution and the Due Process Clause of the Fourteenth Amendment to the U.S. Constitution. 12 The latter, they said, worked in Ohio was almost always practicing Ohio law.”10 But that age, they concluded, has passed. They pointed out that many people keep secondary offices in their homes (or vacation homes) that may be located in Ohio.11

The concurring justices concluded that, as applied to an attorney who is not practicing in Ohio courts or providing Ohio legal services, the Ohio rule on unauthorized practice violates the Ohio Constitution and the Due Process Clause of the Fourteenth Amendment to the U.S. Constitution.12 The latter, they said, protects an individual’s right “to pursue and continue in a chosen occupation free from unreasonable government interference,” and the Ohio rule violated that limitation on state regulation because “it does not bear a rational relationship to any discernable state interest.”13

II. NATIONAL INSTITUTE OF FAMILY AND LIFE ADVOCATES v. BECERRA

Towards the end of the October 2017 Term, on June 26, 2018, the Supreme Court issued its opinion in National Institute of Family and Life Advocates v. Becerra (NIFLA).14 In NIFLA, in relevant part, the Court considered a California law that required certain licensed facilities offering family-planning or pregnancy services to provide their patients with disclosures about the availability of other pregnancy-related services, including abortion. The plaintiffs provided pregnancy-related services but were opposed to abortion. The Court concluded that the law was likely unconstitutional as applied to the plaintiffs, and that, accordingly, the Ninth Circuit’s judgment denying a preliminary injunction against its application to them should be vacated.15

The Court, in an opinion by Justice Clarence Thomas, characterized the law as a content-based regulation of speech that would usually be subject to strict scrutiny, but it discussed whether a lower level of scrutiny should be applied to the law because it was a regulation of “professional speech.”16 The Court held that it had not “recognized ‘professional speech’ as a separate category of speech,”17 although it did not foreclose the possibility that a reason might exist for treating content-based regulations of professional speech differently from content-based regulations of other kinds of speech.18

The Court noted that it had afforded less protection for what might be called professional speech in two circumstances, but that neither of those decisions turned on the fact that professionals were speaking.19 First, the Court had applied more deferential review to laws requiring professionals to disclose factual, noncontroversial information in their “commercial speech.”20 Second, it had permitted states to regulate professional conduct even where that conduct “incidentally involves speech.”21 The NIFLA Court quickly dismissed the first category as irrelevant because the law at issue did not regulate commercial speech about the services that the licensed provider itself offered.22 The Court described the second category as involving regulation of commercial activity that “incidentally” involves speech, and identified “longstanding torts for professional malpractice” as exemplars for such laws. The case it relied most heavily upon to illustrate the category, though, was Planned Parenthood of Southeastern Pennsylvania v. Casey.24 In Casey, the Court upheld a requirement that physicians about to perform abortions obtain “informed consent” by informing their patients about the nature of the abortion procedure, the health risks of abortion and childbirth, and the probable gestational

7 Id. at *6 (DeWine, J., concurring).
8 Id. at *8 (DeWine, J., concurring).
9 Id. at *9 (DeWine, J., concurring). The concurring justices concluded that Jones and others in her position “are not . . . holding themselves out as lawyers to the Ohio public.” Id. Implicit in that finding is that Jones either did not use letterhead with an Ohio address or used such letterhead but with a disclaimer that she did not practice in Ohio or that the use of such a letterhead, even without a disclaimer, would not constitute holding herself out as a lawyer. As noted below, D.C. Rule 49 considers the maintenance of an office in the District of Columbia under similar circumstances (say, the practice of law in Virginia courts) as unauthorized practice in D.C. regardless of any disclaimers on letterhead or other business documents.
11 Id.
12 Id.
13 Id. at *8, *9 (DeWine, J., concurring).
15 Id. at 2376.
16 Id. at 2371-76.
17 Id. at 2371. See also id. at 2372 (“This Court’s precedents do not recognize such a tradition for a category called ‘professional speech.’”).
18 Id. at 2375.
19 Id. at 2372.
20 Id.
21 Id.
22 Id. at 2372.
23 Id. at 2373.
age of the unborn child. The NIFLA Court concluded that the California law before it was not an informed consent law like the one in Casey because it required the provision of information about other services, not those offered by the licensed provider. The Court concluded that these precedents did not establish a professional speech doctrine that would subject content-based laws like California’s to a lower burden of scrutiny based on the fact that the law regulated the speech of professionals.

The notion that regulation of professional speech is treated with greater deference under the First Amendment than regulations of other kinds of speech traces back to Justice Byron White’s concurring opinion in Lowe v. Securities and Exchange Commission. Justice White believed that a professional’s speech would always be only “incidental” to his or her conduct in providing professional services, and that therefore a regulation of that speech would have only an “incidental impact on speech.” He emphasized that the regulation of professional speech would be limited to speech where an attorney-client or similar relationship exists. Prior to NIFLA, a number of lower courts had relied upon Justice White’s Lowe concurrence to conclude that the regulation of professional speech was not subject to heightened First Amendment scrutiny. The validity of those cases, and the authority of Justice White’s Lowe concurrence, is now in doubt.

In NIFLA, the Court expressed great concerns about the application of a “professional speech” doctrine, noting that the category itself was not well-defined, that lower courts had applied the doctrine to any speech involving personalized services and a state license, and that constitutional deference to regulations of such speech “gives the States unfettered power to reduce a group’s First Amendment rights by simply imposing a licensing requirement.”

The Court chose not to determine whether intermediate or strict scrutiny applied to the California law because it concluded that it could not survive even the former. California asserted that its interest was in providing low-income women with information about state-sponsored services, but the Court concluded that the law had too many exceptions of various kinds to be sufficiently drawn to achieve that goal, and that California had not demonstrated that a public-relations campaign could not achieve the interest.

III. Unauthorized Practice of Law Rules Revisited

Rule 49 of the Rules of the District of Columbia Court of Appeals provides that “no person may engage in the practice of law in the District of Columbia . . . unless enrolled as an active member of the D.C. Bar.” The “practice of law” is “providing professional legal advice or services where there is a client relationship of trust or reliance.” “One is presumed to be practicing law” when one does various enumerated acts, including “preparing or expressing legal opinions.” The rule defines “[i]n the District of Columbia” as “conduct in, or conduct from an office or location within, the District of Columbia.”

The Commentary to Rule 49 identifies four “general purposes” for the rule: (1) to protect members of the public from unqualified representation, (2) to ensure that people who hold themselves out to perform or perform the services of lawyers are subject to the D.C. Bar’s disciplinary system, (3) to maintain the efficiency and integrity of the administration of justice and the system of regulation of practicing lawyers, and (4) to ensure that the activities of the D.C. Bar (including its system of regulating lawyers) is supported financially by those exercising the privilege of membership in the D.C. Bar.

Various activities, listed in Rule 49(c), are “permitted as exceptions” to the general rule against the practice of law by individuals who are not active members of the D.C. Bar. Among these are providing legal services before special courts or agencies of the U.S. or D.C., providing legal services related to proceedings in federal court, and providing legal services in D.C. on an incidental or temporary basis. If the person has an office in the District of Columbia, these activities are generally permitted only if the person gives prominent notice of the limitations of his or her practice in all business documents.

There is no exception, however, for an individual who simply practices in another state from an office in the District of Columbia. Regardless of the disclaimers made on business

25 NIFLA, 138 S. Ct. at 2373 (quoting Casey, 505 U.S. at 881).
26 Id. at 2373-74.
28 Id. at 232. Justice White analogized the regulation of professional speech to the regulation of words of “offer and acceptance” under contract law. Id.
29 Id.
30 Serafine v. Branaman, 810 F.3d 354, 359 & n.2 (5th Cir. 2016) (“The Supreme Court has never formally endorsed the professional speech doctrine, though some circuits have embraced it based on Justice White’s concurrence in [Lowe].”). Cf. NIFLA, 138 S. Ct. at 2371 (listing two of the three cases cited by the Fifth Circuit’s Serafine decision without mentioning Justice White’s concurrence).
31 Id. at 2375.
documents, an attorney with an office in D.C., who is a member of the bar of Maryland or Virginia (but not D.C.), and whose practice consists exclusively of representing residents of those states in the courts of those states or concerning transactions in those states, is practicing law in the District of Columbia and violating Rule 49(a). Such violations are also likely to be deemed violations of the rules of professional conduct by the state bars where the practitioners are admitted.41

Rule 49 is not unique in this respect. In Attorney Grievance Commission v. Harris-Smith,42 the Court of Appeals of Maryland sanctioned an attorney, suspending her for thirty days, because she had an office in Maryland but was not admitted to the Maryland Bar. It imposed this sanction despite the fact that she limited her practice to federal bankruptcy matters and was admitted in the federal District of Maryland.43 Indeed, Maryland suspended her from practice even though she was not admitted to the Maryland Bar.44

Similarly, admission to practice before the federal district court in the District of Columbia (D.D.C.) is generally limited to members of the D.C. Bar and “attorneys who are active members in good standing of the Bar of any state in which they maintain their principal law office.”45 Ms. Jones from In re Jones, who had an office in Ohio but practiced in Kentucky, would be ineligible for admission. Similar distinctions appear to be made with respect to participation by non-member attorneys in the D.D.C.46

The concurrence from In re Jones and the Supreme Court’s decision in NIFLA are both relevant to unauthorized practice rules like Rule 49, but in different ways. Precedents can offer factual similarities, similar reasoning, or binding authority. In re Jones is strong on the first element and modest on the second; concurrences from state supreme courts, of course, are not at all binding. NIFLA, on the other hand, is as strong as you can get on the third element, but the facts are not directly related to the unauthorized practice of law. It does, however, call into question some of the basic underlying justification for the rules against certain kinds of such unauthorized practice, particularly the provision of advice about legal matters.

A. Do Jurisdictions Have a Legitimate Interest in Regulating Local Lawyers Practicing in Other States?

The Jones concurrence asks: in today’s world, what is the point of precluding people from working remotely?27 The concurring justices could not find a justification that met rational basis scrutiny. What interest does Ohio have, they asked, in preventing someone from using an office in Ohio if her legal practice only affects another state? Neither consumers of legal services in Ohio nor the conduct of Ohio courts was at issue, and the concurring justices could not identify any other interest that Ohio had in regulating conduct like Jones’s.

The Commentary to Rule 49 suggests other purposes.48 For example, it suggests that the D.C. Bar has interests in subjecting those acting (or holding themselves out) as D.C. attorneys to the D.C. Bar disciplinary system, maintaining the system of regulating the practice of lawyers, and ensuring that those exercising the

---

41 Rule 5.5(a) of the ABA Model Rules of Professional Conduct states: “A lawyer shall not practice law in a jurisdiction in violation of the regulation of the legal profession in that jurisdiction, or assist another in doing so.” See also Attorney Grievance Commission of Maryland v. Walker-Turner, 372 Md. 85 (2002) (holding that a member of the bar of Maryland violated Maryland Rule of Professional Conduct 5.5(a) by practicing from offices in the District of Columbia in violation of Rule 49, and suspending that attorney for thirty days). Rule 49 provides that the Committee on Unauthorized Practice in the District of Columbia may “refer cases” to “other appropriate authorities,” presumably including agencies supervising the conduct of bar members in other states. D.C. Ct. App. R. 49(d)(12)(D).

42 356 Md. 72 (1999).

43 The Court concluded that Harris-Smith’s practice was not strictly limited to bankruptcy matters because she had introductory meetings with clients in which she analyzed their legal problems. Id. at 83-84. Compare Barker, supra note 40, at 1539-40 (arguing that prescreening clients by an attorney authorized to appear before a federal agency would be analogous to a federal court determining its own jurisdiction, and should be permitted); id. at 1545 (“[A] federally admitted lawyer should have the right to evaluate a prospective client’s problem to determine whether federal jurisdiction can be invoked in a court where the lawyer is admitted and, if so, whether other courses, e.g., arbitration, state-court litigation, are preferable.”).

44 Harris-Smith, 356 Md. at 92.

45 D.D.C. Loc. R. 83.8. It is unclear why the rule uses the word “any” (as opposed to “the”) since an attorney presumably will not have more than one “principal law office.”

46 D.D.C. Loc. R. 83.2(c) (providing that non-members of the bar may submit papers joined by member of the D.D.C. bar, but excepting non-members “who engage[] in the practice of law from an office located in the District of Columbia.”). Attorneys with an office in the District of Columbia must be members of the D.C. and the federal court bars. Thus, members of the D.C. bar who work in D.C. (but are not D.D.C. members) cannot file papers joined with federal bar members in the D.D.C. So, too, those with D.C. offices who are not D.C. bar members, but whose practice falls under one of the exceptions in Rule 49(c).

Curiously, the local rules do not prohibit such individuals from being admitted pro hac vice (as opposed to merely filing papers without such admission, which is the focus of Local Rule 83.2(c)). The local rules require that a motion seeking pro hac vice admission for an attorney must state in the motion whether the attorney is a member of the D.C. bar or has an application pending, D.D.C. Loc. R. 83.2(d). But the rule does not require denial of the motion if the applicant is not a member or potential member.

47 Jones, 2018 WL 5076017, at *5 (DeWine, J., concurring) (“We might say—as we do of many employees in today’s world—that [Jones] was working remotely.”).

48 See supra note 37 and accompanying text.
privileges of membership in the D.C. Bar support financially the activities of the D.C. Bar.49 But these purposes simply assume the conclusion that the D.C. Bar has an interest in attorneys with offices in the District regardless of what those attorneys do. The stated purposes do not answer the question why discipline should not be meted out by, and financial support for a bar given to, the state in which the attorneys’ practice is focused, rather than where their offices are located.

In 2017, the D.C. Circuit discerned another interest underlying unauthorized practice rules, at least for those who appear in federal court. In National Association for the Advancement of Multijurisdiction Practice v. Howell (NAAMJP),50 the plaintiff challenged the “principal office” rule of the D.D.C., which (in relevant part) limits bar membership to D.C. Bar members and those who are active members of the bar of the state in which they have their principal office.31 This “Principal Office Provision” was rational, the Court held, because the federal district court had an interest in making sure that attorneys who practice in its court “are subject to supervision by the state to which their practice is most geographically proximate.”52 With respect to Alice Jones, for example, the court would say that the D.D.C. has an interest in making sure that Ohio, and not Kentucky, could supervise Ms. Jones’s practice. Why exactly a federal district court would care about which other courts could discipline an attorney for other conduct—a federal district court, of course, can adequately police the conduct of attorneys in its own court—was not exactly clear.53 Nor is it clear why that federal court would prefer the “geographically proximate” jurisdiction over the one where the attorney performs all of his or her services for citizens of that jurisdiction. It is especially befuddling because the rule authorizes admission of D.C. Bar members regardless of where they live or work, so it assumes a bar can supervise attorneys with offices outside its jurisdiction.

In any event, an interest sufficient to justify denial of admission to a specific court might not be sufficient to justify precluding someone from providing legal advice or doing anything that might constitute the practice of law throughout a jurisdiction.54 The In re Jones concurring justices, after all, relied on the right to pursue a chosen profession to conclude that Ohio’s unauthorized practice rule was unconstitutional.55 Thus, even if the D.C. Circuit’s reasoning were convincing, it is not clear that it would justify precluding someone from being admitted into a state bar.

The concurring justices in In re Jones have a strong case. A state should not be able to force an attorney to join its bar and pay dues just because that attorney works out of some location in that jurisdiction—whether an office, a home, or even a public place—if the attorney has no other contact with the jurisdiction.56

B. Do Regulations of the Provision of Legal Advice Have to Satisfy Strict Scrutiny?

NIFLA may present the more serious threat to unauthorized practice rules because it questions the very premise of the application of such rules to a central part of the practice of law: the provision of legal advice. The premise of unauthorized practice rules is that providing advice about legal matters to an individual is:

1. conduct and not speech, or at least that the speech element is just incidental to the conduct, or

2. professional speech, where judicial review under the First Amendment is more deferential.57

NIFLA casts grave doubt on both of these possibilities. It notes that certain lower courts have recognized professional speech as a separate category, defining it as speech by individuals who provide personalized services to clients based on their expert knowledge and judgment and who are subject to a generally applicable licensing scheme.58 But the Court asserts that any lower level of scrutiny for professional speech in its own cases had nothing to do with the fact that professionals were speaking. Rather, although the Court has upheld regulations of professional conduct that incidentally involved speech, it does not automatically assume that regulations that apply to professionals are always regulations of conduct. In NIFLA itself, the plaintiffs provided advice on pregnancy-related procedures and conducted some procedures themselves, and the law required them to provide additional information on other procedures (including abortion) that were available in California. This notice requirement, the Court held, “is not a[] . . . regulation of professional conduct.”59 Rather, it “regulates speech as speech.”60 The mere fact that the plaintiffs were licensed professionals did not render all of their advice regulable conduct.

---

49 Id.
50 851 F.3d 12 (D.C. Cir. 2017).
52 NAAMJP, 851 F.3d at 18.
53 Kentucky apparently has no requirement that a member of the bar either live or work in Kentucky. Compare Schoenfeld v. Schneiderman, 821 F.3d 273 (2d Cir. 2016) (upholding NY Jud. Law § 470, which requires non-resident bar members to have an office in New York).
54 Compare Leis v. Flynt, 439 U.S. 438 (1979) (holding that there was no Fourteenth Amendment property interest in being admitted pro hac vice).
55 Jones, 2018 WL 5076017, at *8 (DeWine, J., concurring) (“The Fourteenth Amendment to the federal Constitution also has been held to protect the right of an individual to pursue and continue in a chosen occupation free from unreasonable government interference.”).
56 D.C. Ct. App. R. 49 states that “no person may engage in the practice of law in the District of Columbia . . . unless enrolled as an active member of the D.C. Bar” and defines “in the District of Columbia” as “conduct in, or conduct from an office or location within, the District of Columbia,” D.C. Ct. App. R. 49(a), 49(b)(3). Thus, unless it falls within the “incidental and temporary practice” exception of D.C. Ct. App. R. 49(c)(13), work out of a home (or a library or a restaurant) in D.C. constitutes the practice of law in D.C.
58 NIFLA, 138 S. Ct. at 2371. 59 Id. at 2373.
60 Id. at 2374.
It would be hard to distinguish the pregnancy-related advice and services at issue in NIFLA from legal advice and services regulated under Rule 49 and other unauthorized practice prohibitions. Legal advice is presented in the form of speech; there is no more reason to treat it as conduct than the pregnancy-related advice at issue in NIFLA. Moreover, a prohibition on “legal advice” is a regulation based on content: we determine whether the prohibition has been violated by looking at the content of the speech and determining whether it is legal advice. And while the state might have an important interest in protecting consumers from advice that does not meet professional standards under certain circumstances, it may be difficult to defend a total prohibition on legal advice from particular individuals as the most narrowly-tailored means of protecting that interest. A tort remedy for malpractice or a requirement of full disclosure of qualifications would more precisely protect that interest without unnecessarily restricting non-harmful speech.

Indeed, the rules of professional conduct specifically permit licensed attorneys to take representations in subject matters where they may not yet be particularly learned.61 Hence, the protection of consumers from incompetent legal advice cannot justify a rule precluding a lawyer from providing legal advice to clients in a state where the lawyer is admitted from an office in a jurisdiction where he is not. Indeed, it is hard to justify a rule prohibiting that attorney from providing legal advice to individuals in the jurisdiction where he is not admitted about the law where he is admitted, or even counseling individuals about the law in a jurisdiction where he is not admitted, provided that proper disclosures are made about the attorney’s qualifications. Advice, NIFLA tells us, is speech, and content-based prohibitions on speech must meet heightened scrutiny.

IV. Concluding Thoughts

Reading through the rules regarding the scope of unauthorized practice and the justifications used to support them, one cannot help but get the feeling that they owe their continued existence to tradition and inertia, and that they have not kept up with the times.62 What could possibly justify a rule prohibiting an attorney admitted in Maryland from working out of a home office in D.C. to represent clients in Maryland? In the age of electronic communication, why would a state need to have attorneys admitted to its bar maintain a physical office in its jurisdiction in order to serve them with process?63 A cynic might think that the rules primarily serve the purposes of restricting competition in the legal profession and lining the coffers of bar associations.64 NIFLA and Jones have shed some light (or, perhaps, shadows) on the basic principles upholding restrictions on legal practice. It might be time for us to reconsider these restrictions in light of them.

61 See D.C. Rule of Prof. Cond. 1.1, cmt. 2 (“A lawyer can provide adequate representation in a wholly novel field through necessary study.”). See also David McGowan, Two Ironies of UPL Laws, 20 CHAPMAN L. REV. 225, 226 (2017) (“Passing the bar exam does not entail practical competence in any particular field.”); id. at 244 (“A licensed lawyer who accepts a matter and plans to learn by doing does not violate any UPL restriction, while an experienced legal assistant who is competent to handle the matter would.”).

62 Jones, 2018 WL 5076017, at *9 (DeWine, J., concurring) (“In an earlier age, perhaps such a rule made sense. Before the advent of the Internet, electronic communication, and the like, a lawyer who worked in Ohio was almost always practicing Ohio law.”).

63 See supra note 53.

The Supreme Court’s unanimous 2012 decision in *Hosanna-Tabor Evangelical Lutheran Church and School v. EEOC* firmly established what the federal courts of appeals had previously recognized for decades: that religious ministries have an absolute First Amendment right to select their own religious ministers, free from government interference.¹ The Court, like many courts of appeals before it, explained that the founders protected this right through both Religion Clauses of the First Amendment. By “forbidding the ‘establishment of religion’ and guaranteeing the ‘free exercise thereof,’ the Religion Clauses ensured that the new Federal Government—unlike the English Crown—would have no role in filling ecclesiastical offices.”²

But as the Court’s first foray into defining and applying the “ministerial exception” doctrine, *Hosanna-Tabor* understandably did not answer all questions about how the doctrine operates. Some of those questions are important. Such as how to determine what a religious ministry is, who a religious minister is, what types of government interference are impermissible, and how a substantive right grounded in both Religion Clauses should operate at a procedural level. To give concrete examples: does a Jewish day school count as a ministry, even if it has an equal opportunity policy that forbids religious discrimination in employment, receives government funding, and accepts non-Jewish students? Is the principal of a Catholic elementary school a minister, even if she has neither formal religious training nor an explicitly religious title? Is enforcing a ministerial contract’s for-cause termination provision impermissible interference, even where the ministry’s basis for termination was not the quality of the minister’s sermons? Finally, can the ministerial exception be lost in whole or in part via procedural means, such as waiver or inability to raise it immediately on interlocutory appeal?

To be sure, the Court’s analysis in *Hosanna-Tabor* informs, and in some instances largely answers, these questions. But lower courts are now working through the answers to all of them. This article provides a survey of what conclusions the courts are reaching.

I. The Ministerial Exception

The First Amendment to the United States Constitution guarantees individuals the right to the free exercise of religion and prohibits the establishment of religion by the federal government.³ Through the doctrine of incorporation, the Free Exercise and Establishment Clauses have been applied to the

---

² Id. at 184.
³ U.S. Const. amend. I (“Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances.”).
forms the Due Process Clause of the Fourteenth Amendment. The Supreme Court has long recognized that both Religion Clauses together “radiate[] . . . a spirit of freedom for religious organizations, an independence from secular control or manipulation” that places “matters of church government and administration beyond the purview of civil authorities.” Lower courts have also long recognized that this broad principle of religious autonomy includes a subset of specific protections regarding employment decisions made by religious organizations with respect to employees who serve in religiously significant roles: the so-called “ministerial exception.” In 2012, in the context of an employment discrimination case, the U.S. Supreme Court formally agreed.

II. *Hosanna-Tabor v. EEOC*

Hosanna-Tabor Evangelical Lutheran Church and School operated an elementary school in Redford, Michigan, “offering a ‘Christ-centered education’” to its students. Hosanna-Tabor is a congregation of the Lutheran Church-Missouri Synod, and it employs both “called” and “lay” teachers to educate the children at its school. Called teachers are regarded as having been called to their vocation by God through a congregation, and they must satisfy certain academic requirements which may include taking courses in theology, obtaining an endorsement from a local synod, and passing an oral examination. A qualified teacher may receive a religious “calling” by a congregation, which entitles the teacher to receive the formal title “Minister of Religion, Commissioned.” A called minister serves for an “open-ended term” and can only be terminated “for cause and by a supermajority vote of the congregation.” By contrast, lay teachers are appointed by the school board for one-year renewable terms and do not have to be Lutheran or “trained by the Synod.” Called and lay teachers “generally perform[] the same duties,” but lay teachers are only hired when called teachers are not available.

In 1999, Cheryl Perich was hired as a lay teacher; she later became a called teacher after she satisfied the requirements and received a “diploma of vocation” designating her a commissioned minister.” Perich initially taught kindergarten, but later taught fourth grade. In addition to several secular subjects, Perich taught a religion class, led students in forty-five minutes of prayer and devotional exercises each day, and brought her students to a weekly chapel service which she led twice a year.

Perich was diagnosed with narcolepsy after presenting symptoms of “sudden and deep sleeps from which she could not be roused.” She took disability leave beginning in her sixth year of teaching, but notified the school that she would be able to return to work in February of 2005. Perich was advised that a lay teacher had been retained to fill the position for the remainder of the year, and concern was expressed that she “was not yet ready to return to the classroom.” Shortly after Perich notified the school that she was ready to return to work, the congregation met and voted to offer Perich a “peaceful release” from her ministerial calling, offering to pay a portion of her health insurance premiums in return for her resignation. Perich refused to resign and refused to participate in the internal dispute resolution required under the school’s Lutheran beliefs, so her employment was terminated.

---

4 See Cantwell v. Connecticut, 310 U.S. 296 (1940) (incorporating free exercise clause); Everson v. Bd. of Educ., 330 U.S. 1 (1947) (incorporating establishment clause). See also U.S. Const. amend. XIV, § 1 (“All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the State (or States) in which they reside. No State shall . . . abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.”). On the doctrine of incorporation, see generally Jerold H. Israel, *Selective Incorporation Revisited*, 71 Geo. L. J. 253 (1982). Justice Clarence Thomas has questioned whether the Establishment Clause should have been incorporated, as he sees it as a rule against federal establishments and not state establishments. See, e.g., *Elk Grove Unified Sch. Dist. v. Newdow*, 542 U.S. 1 (2004) (Thomas, J., concurring in the judgment). It is notable, though, that he joined the Court’s opinion in *Hosanna-Tabor*, which disposed of both federal- and state-law claims under the Establishment Clause. Justice Thomas wrote separately only to address how courts should define ministerial status, and he did not address the question of whether the Establishment Clause should have applied to the state-law claims at issue. Of course, the plaintiff had conceded that her state-law claims rose or fell based upon her federal-law claim, *Hosanna-Tabor*, 565 U.S. at 194 n.3, and both sets of claims would have lost under the Free Exercise Clause in any event. So Justice Thomas may not have felt the need to clarify his view on the issue. Another possibility is that his view could leave room for the Establishment Clause acting as a structural barrier to both federal- and state-law claims that would entangle the government in internal religious affairs. See *Conlon v. InterVarsity Christian Fellowship*, 777 F.3d 829 (6th Cir. 2015) and *Lee v. Sixth Mount Zion Baptist Church*, 903 F.3d 113 (3d Cir. 2018).


6 *See, e.g., McClure*, 460 E.2d at 559. The principle of religious autonomy comes up in several related and overlapping contexts, including the church-autonomy doctrine and the ecclesiastical abstention doctrine. *See, e.g., Korte v. Sebelius*, 735 F.3d 654, 677 (7th Cir. 2013) (church-autonomy doctrine); *Puri v. Khalsa*, 844 F.3d 1152, 1164 (9th Cir. 2017) (ecclesiastical abstention doctrine).

7 *Hosanna-Tabor*, 565 U.S. at 177.
Perich filed a charge of disability discrimination with the EEOC. The EEOC brought suit against Hosanna-Tabor under the Americans with Disabilities Act and the Michigan Persons with Disabilities Civil Rights Act, and Perich intervened in the suit. Hosanna-Tabor moved for summary judgment under the First Amendment’s ministerial exception, arguing that Perich was a minister and that she was terminated by a religious organization for religious reasons. The district court agreed and granted summary judgment in Hosanna-Tabor’s favor. The U.S. Court of Appeals for the Sixth Circuit reversed, finding that Perich did not qualify as a “minister” under the exception because her religious duties as a called teacher were the same as the duties of lay teachers, and in any event only consumed forty-five minutes of each school day.

A. The Supreme Court’s Opinion

In a unanimous decision written by Chief Justice John Roberts, the U.S. Supreme Court reversed the Sixth Circuit and found that the ministerial exception precluded Perich from pursuing her employment claims against Hosanna-Tabor. The Court held that both the Free Exercise and Establishment Clauses of the First Amendment “bar the government from interfering with the decision of a religious group to fire one of its ministers.”

The Court began its analysis by determining the meaning of the Religion Clauses through a detailed examination of the history of tensions between church and the state, from the time of the Magna Carta and through the time of the founding of the United States. The Court ruled that “[i]t was against this background of church-state conflict “that the First Amendment was adopted,” and that part of the founders’ purpose was to prevent the kind of state-sanctioned ministerial selection that had created so much conflict in England and its colonies. The Court then explained how this principle of religious autonomy was reflected in early Supreme Court decisions addressing property disputes between religious entities. Finally, the Court acknowledged the “extensive experience” that the federal courts of appeals had obtained in administering the ministerial exception over previous decades, and how the lower courts had “uniformly recognized” that the First Amendment required a ministerial exception to certain state and federal employment claims brought by ministers against religious organizations.

After recognizing the ministerial exception’s existence, the Court turned to its application. There was no dispute that Hosanna-Tabor was the kind of entity that could assert the ministerial exception. It was also undisputed that nondiscrimination laws could be applied to inappropriately interfere with internal church affairs. So the primary question before the Court was whether Perich held a ministerial role for the school.

The Court noted that the courts of appeals were in agreement that the ministerial exception was “not limited to the head of a religious congregation,” and so that Perich was a teacher instead of a pastor was not dispositive. But the Court declined to “adopt a rigid formula” to determine ministerial status in its “first case involving the ministerial exception.” Instead, the Court concluded that the facts before it were sufficient to find that Perich was a minister. The Court identified four “considerations” supporting its conclusion: Perich’s (1) “formal title,” (2) “the substance reflected in that title,” (3) her “use of th[e] title,” and (4) “the important religious functions she performed.”

The Court recounted that: Perich taught her students religion four days a week, and led them in prayer three times a day. Once a week, she took her students to a school-wide chapel service, and—about twice a year—she took her turn leading it, choosing the liturgy, selecting the hymns, and delivering a short message based on verses from the Bible. During her last year of teaching, Perich also led her fourth graders in a brief devotional exercise each morning. As a source of religious instruction, Perich performed an important role in transmitting the Lutheran faith to the next generation.

22 Id. at 180-81.
23 Id. at 181.
24 Id. at 195.
25 Id. at 181.
26 Id. at 182-86.
27 Id. at 185-87. See Watston, 13 Wall. at 727 (“[W]henever the questions of discipline, or of faith, or ecclesiastical rule, custom, or law have been decided by the highest of [the] church judicatories to which the matter has been carried, the legal tribunals must accept such decisions as final, and as binding on them.”); Redoff, 344 U.S. at 116 (“[O]ur opinion in Watson ] “radiates . . . a spirit of freedom for religious organizations, an independence from secular control or manipulation—in short, power to decide for themselves, free from state interference, matters of church government as well as those of faith and doctrine”); Serbian Eastern Orthodox Diocese v. Milivojevich, 426 U.S. 696, 724 (1976) (the First Amendment “permit[s] hierarchical religious organizations to establish their own rules and regulations for internal discipline and government, and to create tribunals for adjudicating disputes over these matters,” free from judicial gainsaying).
“In light of these considerations,” the Court concluded that “Perich was a minister covered by the ministerial exception.” 35

The Court then identified three ways that the lower court had gone astray. First, it failed to consider Perich’s title, which was relevant given what it signified: “that an employee has been ordained or commissioned as a minister,” and “that significant religious training and a recognized religious mission underlie the description of the employee’s position.” 36 Second, “the Sixth Circuit gave too much weight to the fact that lay teachers at the school performed the same religious duties as Perich.” 37 Finally, “the Sixth Circuit placed too much emphasis” on Perich’s “performance of secular duties” and the fact that “her religious duties consumed only 45 minutes of each workday.” 38 Noting and rejecting the EEOC’s “extreme position” that the ministerial exception should cover only those who “perform exclusively religious functions,” the Court explained that even the “heads of congregations themselves often have a mix of duties” both sacred and secular, and thus that ministerial status cannot “be resolved by a stopwatch.” 39 The Court emphasized that the proper touchstone was “the nature of the religious functions performed,” along with the three “other considerations” it had identified. 40

The Court also rejected Perich’s and the EEOC’s arguments that they should at least be permitted to probe whether the religious reasons for her firing were “pretextual.” That approach, the Court explained, “misses the point”: the ministerial exception should cover only those who “perform exclusively religious functions,” the Court explained that even the “heads of congregations themselves often have a mix of duties” both sacred and secular, and thus that ministerial status cannot “be resolved by a stopwatch.” 39 The Court emphasized that the proper touchstone was “the nature of the religious functions performed,” along with the three “other considerations” it had identified. 40

The Court also rejected Perich’s and the EEOC’s arguments that they should at least be permitted to probe whether the religious reasons for her firing were “pretextual.” That approach, the Court explained, “misses the point”: the ministerial exception does not “safeguard a church’s decision to fire a minister only if secular courts could second-guess the organization’s sincere determination that a given employee is a ‘minister’ under the organization’s theological tenets.” 41

In conclusion, the Court acknowledged society’s important interest in employment nondiscrimination, but found that the “First Amendment has struck the balance” in favor of religious autonomy: “church[es] must be free to choose those who will . . . preach their beliefs, teach their faith, and carry out their mission.” 42 The Court left to another day how the First Amendment required weighing claims that sounded in contract or tort instead of nondiscrimination statutes. 43

B. The Two Concurring Opinions

Justice Clarence Thomas filed a concurring opinion on his own behalf, and Justice Samuel Alito filed one joined by Justice Elena Kagan. Both concurring opinions focused on the question of how to define ministerial status.

Justice Thomas argued that the standard should be one that “defer[s] to a religious organization’s good-faith understanding of who qualifies as its minister.” 44 He reasoned that “a religious organization’s right to choose its ministers would be hollow . . . if secular courts could second-guess the organization’s sincere determination that a given employee is a ‘minister’ under the organization’s theological tenets.” 45

Justices Alito and Kagan warned against overreading the Court’s four considerations, explaining that “it would be a mistake if the term ‘minister’ or the concept of ordination were viewed as central to the important issue of religious autonomy that is presented in cases like this one.” 46 They explained that the Court’s unanimous decision was consistent with a preexisting “functional consensus” among the lower courts that the focus of ministerial analysis should be “on the function performed by persons who work for religious bodies.” 47 The Justices recounted how, in the four decades of ministerial exception caselaw, the overwhelming majority of circuits and state supreme courts “ha[d] concluded that the focus should be on the function of the position” in “evaluating whether a particular employee is subject to the ministerial exception.” 48 They accordingly reasoned that the ministerial exception “should apply to any employee who leads a religious organization, conducts worship services or important religious ceremonies or rituals, or serves as a messenger or teacher of its faith.” 49 They warned that failing to adopt this approach and overemphasizing ministerial titles or ordination would necessarily leave members of non-Protestant faith groups—such as “Catholics, Jews, Muslims, Hindus, or Buddhists,” who do not use always such titles or have different ways of commissioning religious leaders—unprotected by the First Amendment. 50

III. APPLYING Hosanna-Tabor IN THE LOWER COURTS

Courts applying Hosanna-Tabor have asked four main questions: (1) what is a “ministry,” (2) who is a “minister,” (3) what counts as impermissible interference, and (4) how does the ministerial exception operate procedurally?

A. What Is a “Ministry”?

In Hosanna-Tabor, there was no question that the petitioners—a Lutheran church and a Lutheran elementary school—were “ministries” for purposes of the ministerial exception. But several cases have since raised that question, and courts have answered it in two ways.

Some courts have looked at the religious nature of the party asserting the ministerial exception as a whole. For instance, the

---

35 Id.
36 Id. at 193.
37 Id.
38 Id.
39 Id. at 193-94.
40 Id. at 194.
41 Id. at 194-95 (internal citation omitted).
42 Id. at 196.
43 Id.
44 Id. (Thomas, J., concurring).
45 Id. at 197.
46 Id. at 198 (Alito, J., concurring).
47 Id. at 198, 204.
48 Petruska v. Gannon Univ., 462 F.3d 294, 304 n.6 (3d Cir. 2006) (collecting cases from the D.C., Fourth, Fifth, and Seventh Circuits) (internal quotations omitted).
49 Hosanna-Tabor, 565 U.S. at 199.
50 Id. at 198.
first time the issue arose after *Hosanna-Tabor* was when the Sixth Circuit faced the question of whether a national parachurch organization that serves on college campuses, InterVarsity Christian Fellowship, was a "ministry." In *Conlon v. InterVarsity Christian Fellowship/USA*, the court rejected the idea that the doctrine applies only to houses of worship such as churches, synagogues, and mosques.\(^{51}\) Instead, relying on the Fourth Circuit’s 2004 decision in *Shalliehsabou v. Hebrew Home of Greater Washington, Inc.*, the court concluded that a group is a religious organization for purposes of the ministerial exception if its "mission is marked by clear or obvious religious characteristics."\(^{52}\) Under this rule, InterVarsity was a "ministry" because its avowed public purpose was "to advance the understanding and practice of Christianity in colleges and universities."\(^{53}\)

The Seventh Circuit’s *Grussgott v. Milwaukee Jewish Day School, Inc.* decision likewise adopted *Hebrew Home* in 2018, agreeing that the "key inquiry" was whether the institution had a "religious character."\(^{54}\) Further, the court found that a religious institution does not lose its religious identity either by failing to participate in an "ecclesiastical hierarchy" or by having a nondiscrimination policy that allows members of other faiths to receive services or employment from the institution. The Seventh Circuit concluded that imposing such limitations would both interfere in internal religious affairs and discriminate against religious groups that have either less hierarchical structures or more ecumenical ministries.

Courts have also used the *Hebrew Home* test to find that an organization is not eligible to assert the ministerial exception. The Sixth Circuit’s 2018 decision in *EEOC v. R.G. & G.R. Harris Funeral Homes, Inc.* found that a funeral home lacked the requisite religious character to invoke the ministerial exception because, among other things, it had "virtually no religious characteristics," did not seek to "establish or advance" its religious beliefs, and did not "avow any religious purpose" in its articles of incorporation.\(^{55}\)

The second approach asks *Hebrew Home*'s "religious character" question, but it doesn’t look at the employer as a whole. Rather, it looks more narrowly at the nature of the employer at the point of its employment relationship with the plaintiff minister. Thus, in *Penn v. New York Methodist Hospital* and *Scharon v. St. Luke’s Episcopal Presbyterian Hospital*, the Second and Eighth Circuits, respectively, addressed employers with religious heritage that had arguably waned in influence over time. Indeed, in *Scharon*, the employer hospital had become "primarily a secular institution."\(^{57}\) But in both cases, instead of considering whether the employer as a whole qualified as a "ministry," the courts examined the employer’s specific relationship with the suing employee—in both cases, a chaplain. In *Penn*, the court found that the hospital’s Department of Pastoral Care was "marked by clear or obvious religious characteristics," and that this was enough to warrant application of the ministerial exception to the Department’s relationship with its chaplain.\(^{58}\)

**B. Who Is a "Minister"?**

Since the Court’s opinion in *Hosanna-Tabor*, several federal courts of appeals and two state supreme courts have squarely addressed the definition of "minister." Despite the Supreme Court’s repeated assurance that it was not creating a rigid test, the courts have sometimes struggled analytically to determine what to do with the Supreme Court’s four "considerations" for determining ministerial status—title, substance of title, use of title, and function of position. No courts believe all four considerations are necessary; several have found that showing one or two is sufficient where function is among the considerations shown. Only one court has held that a showing of religious function *must* be accompanied by a showing of another consideration. And all of the courts, save the one, have agreed that the "functional consensus" identified in Justice Alito’s concurrence is the touchstone for analyzing whether someone is a minister.\(^{59}\)

The first post-*Hosanna-Tabor* federal appellate case to consider this question, the Fifth Circuit’s *Cannata v. Catholic Diocese of Austin*, leaned heavily on function to determine ministerial status.\(^{60}\) There, the court addressed a music director’s argument that he was not a minister because “he merely played the piano at Mass and . . . his only responsibilities were keeping the books, running the sound system, and doing custodial work, none of which was religious in nature.”\(^{61}\) But the Fifth Circuit found it had "enough" basis to apply the ministerial exception because "there [wa]s no genuine dispute that Cannata played an integral role in the celebration of Mass and that by playing the piano during services, Cannata furthered the mission of the church and helped convey its message to the congregants."\(^{62}\) In other words, “[b]ecause Cannata performed an important function...

---

51 777 F.3d 829 (6th Cir. 2015).
52 Id. at 834 (quoting Shalliehsabou v. Hebrew Home of Greater Wash., Inc., 563 F.3d 299, 310 (4th Cir. 2004)).
53 Id. at 833-34. See also Yin v. Columbia Inst’l Univ., 335 F. Supp. 3d 803, 814-15 (D.S.C. 2018) (applying the Hebrew Home test and determining that a religious university was protected by ministerial exception because the college ‘trains Christians for global missions, full-time vocational Christian ministry in a variety of strategic professions, and marketplace ministry’
54 882 F.3d 655, 658 (7th Cir. 2018).
55 884 F.3d 560, 582 (6th Cir. 2018).
57 Scharon, 929 F.2d at 362.
58 Penn, 884 F.3d at 425.
60 700 F.3d 169, 171 (5th Cir. 2012).
61 Id. at 177.
62 Id. See also Demkovich v. St. Andrew the Apostle Parish, 2017 WL 4393817 *3 (N.D. Ill. 2017) (“By selecting music for mass, Demkovich helped to ‘convey[] the Church’s message’ through the important religious function of worship music.”).
during the service,” he was a minister.63 By contrast, Cannata’s lack of formal religious training was “immaterial.”64 Nor did it matter that he did not hold a formal religious role under church law, since courts “may not second-guess whom the Catholic Church may consider a lay liturgical minister under canon law.”65

About the same time that Cannata came down, the Massachusetts Supreme Judicial Court held that function alone suffices to prove ministerial status.66 In that case, the court held that a teacher at a Jewish school was covered by the ministerial exception even though her role did not obviously meet any of the other three Hosanna-Tabor considerations: “she was not a rabbi, was not called a rabbi, . . . did not hold herself out as a rabbi,” and had not been proven to have received “religious training.”67 The court found it dispositive that “she taught religious subjects at a school that functioned solely as a religious school” for children.68 Two years later, the Kentucky Supreme Court held that, in considering the totality of the circumstances of an employee’s role serving within a religious organization, courts should focus on “actual acts or functions conducted by the employee.”69

Next, the Sixth Circuit’s 2015 Conlon decision concerned an employee who alleged that her termination from the position of “Spiritual Director” violated state and federal employment discrimination law. The court analyzed all four Hosanna-Tabor considerations to determine whether she was a “minister.” First, the court found that the job title of “Spiritual Director” conveyed a religious role.70 On both the substance-of-title and use-of-title considerations, the court found that the plaintiff’s formal title of “Spiritual Director” violated state and federal employment discrimination law. The court analyzed all four Hosanna-Tabor considerations to determine whether she was a “minister.”71

Two years later, in Fratello v. Archdioce of New York, the Second Circuit likewise walked through the four Hosanna-Tabor considerations to determine the ministerial status of a former principal of a Roman Catholic school.72 The court began by acknowledging that “Hosanna-Tabor instructs only as to what we might take into account as relevant, including the four considerations on which it relied; it neither limits the inquiry to those considerations nor requires their application in every case.”73 With that framing, the court proceeded to examine each of the considerations. First, it found that the plaintiff’s formal title of “lay principal” was not sufficiently religious to suggest that the plaintiff performed any religious functions or held a clergy-type role.74 But the court rejected the plaintiff’s argument that this finding was dispositive, since “the substance of the employees’ responsibilities in their positions is far more important.”75 On the substance-of-title consideration, the court found that the plaintiff’s lack of formal religious training or education was not determinative, but rather took a back seat to the fact that the role required her to “be a ‘practicing Catholic in union with Rome’” and “demonstrate proficiency in a number of religious areas” such as “encouraging spiritual growth” and “exercising spiritual leadership” sufficient to “provide ‘Catholic leadership to the School[].’”76 Concerning use-of-title, the court found that the plaintiff knew “that she would be perceived as a religious leader,” and that she held herself out as such through her role in leading “school prayers,” “convey[ing] religious messages in speeches and writings,” and express[ing] the importance of Catholic prayer and spirituality in newsletters to parents.”77 Finally, the court turned to the fourth consideration—religious “functions performed”—and emphasized that performance of “important religious

63  Id. at 178.
64  Id. at 179-80. The Fifth Circuit cited both the majority and concurring opinions in Hosanna-Tabor as support for declining to “second-guess” the church’s decision on who is a lay minister under canon law. Id. at 179.
66  Id. at 486.
67  Id.
68  Id.
69  Kirby v. Lexington Theological Seminary, 426 S.W.3d 597, 613 (Ky. 2014).
70  Conlon, 777 F.3d at 834-35.
71  Id. at 835.
72  Id.
73  Id.
74  863 F.3d 190 (2d Cir. 2017).
75  Id. at 204-05 (emphasis in original).
76  Id. at 206-07.
77  Id. at 207.
78  Id. at 208.
79  Id. See also Ginalski v. Diocese of Gary, 2015 U.S. Dist. LEXIS 168014 at *2 (N.D. Ind. 2016) (district court dismissed employment discrimination claims of terminated Catholic High School Principal finding the ministerial exception applicable). But see Richardson v. Northwest Christian Univ., 242 F. Supp. 3d 1152, 1145-46 (D. Ore. 2017) (district court held that an assistant professor at a private, non-profit, Christian university was not subject to the ministerial exception where plaintiff had a secular job title, had not undergone religious training prior to assuming the position, had not held herself out as a minister, and, while she did perform some “important religious functions” . . . “she was not tasked with performing any religious instruction and she was charged with no religious duties such as taking students to chapel or leading them in prayer”); Bohnert v. Roman Catholic Archbishop of San Francisco, 136 F. Supp. 3d 1094, 1114 (N.D. Cal. 2015) (district court found that a biology teacher at a Catholic all-boys college preparatory school was not subject to the ministerial exception where plaintiff was not an ordained minister nor held out as one by the defendant, had no formal religious or theological studies, and where plaintiff did not “provide spiritual or religious guidance” to students).
functions” was the “most important consideration.”80 Because, “as principal, Fratello ‘conveyed’ the School’s Roman Catholic ‘message and carried out its mission’” the court concluded that “[t]his fundamental consideration therefore weighs strongly” in favor of finding ministerial status, which it did.81

A year later, in early 2018, the Seventh Circuit in Grussgott v. Milwaukee Jewish Day School addressed the ministerial status of a teacher of Hebrew and Jewish studies at a Jewish day school.82 After determining that the school was eligible to raise the ministerial exception, the court turned to the teacher’s ministerial status. Like the courts in Conlon and Fratello, the court walked through the four Hosanna-Tabor considerations.83 It found that the formal title and use-of-title considerations “cut . . . against applying the ministerial exception,” in part because there was “no evidence that Grussgott ever held herself out to the community as an ambassador of the Jewish faith” or otherwise “understood that her role would be perceived as a religious leader.”84 But the court found that both the substance-of-title and function considerations weighed in favor of applying the exception, since her role “entails the teaching of the Jewish religion to students” and since she in fact carried out those religious duties, teaching “about Jewish holidays, prayer, and the weekly Torah readings” and “practic[ing] the religion alongside her students by praying with them and performing certain rituals.”85 While adopting a “totality-of-the-circumstances test,” the court ultimately relied on the Alito concurrence to conclude that “the importance of Grussgott’s role as a ‘teacher of [ ] faith’ to the next generation outweighed other considerations.”86

Later in 2018, in Biel v. St. James School, a split panel of the Ninth Circuit was the first to break with what Justices Alito and Kagan described as the “functional consensus” that function is the focus of ministerial status analysis.87 In Biel, much like in Hosanna-Tabor, the plaintiff was a teacher who taught fifth grade for a religious school and had duties that included teaching a religion class, praying with her students, taking her students to Mass, and embodying the faith to her students. But Biel’s two-judge majority found that the ministerial exception didn’t apply because it believed that the teacher had a less religious title, had received less religious training, and did not hold herself out as a religious leader to the same extent as the fourth grade teacher in Hosanna-Tabor. The panel majority then found that function “alone” cannot determine ministerial status, and questioned whether Grussgott was correctly decided.88 The dissenting judge disagreed, saying that the plaintiff’s duties were “strikingly similar” to those at issue in Hosanna-Tabor, and that the case “is not distinguishable from Grussgott.”89 A petition for en banc rehearing is pending in that case at the time of publication of this article.90

C. What Interference Is Impermissible?

At its core, the ministerial exception protects a religious group’s right under the Free Exercise Clause “to shape its own faith and mission through its appointments,” and it enforces the Establishment Clause’s structural limitation on government “involvement” in “determin[ing] which individuals will minister to the faithful.”91 But what kinds of legal claims violate those core protections?

Employment nondiscrimination claims clearly do.92 And these types of claims are the vast majority of the claims currently brought by ministers against ministries.93 Further, as Hosanna-Tabor explained, the ministerial exception applies to such claims

---

80 Fratello, 863 F.3d at 208-09. Some district courts in the Second Circuit have taken the approach that “the more religious the employer institution is, the less religious the employee’s functions must be to qualify.” Stabler v. Congregation Emanu-El of N.Y., 2017 U.S. Dist. LEXIS 118964, at *18 (S.D.N.Y. 2017) (quoting Penn v. N.Y. Methodist Hosp., 158 F. Supp. 3d 177, 182 (S.D.N.Y. 2016)) (noting that “[t]he ministerial exception should be viewed as a sliding scale, where the nature of the employer and the duties of the employee are both considered in determining whether the exception applies”).

81 Id. at 209.

82 882 F.3d at 656.

83 Id. at 658-59. The Seventh Circuit noted that “other courts of appeals have explained that the same four considerations need not be present in every case involving the exception.” Id. at 658.

84 Id.

85 Id. at 660.

86 Id. at 661 (quoting Hosanna-Tabor, 565 U.S. at 199 (Alito, J., concurring)).

87 911 F.3d 603 (9th Cir. 2018); Hosanna-Tabor, 565 U.S. at 198 (Alito, J., concurring). See also Hollins v. Methodist Healthcare, Inc., 474 F.3d 223, 226 (6th Cir. 2007) (identifying function as “general rule”); EEOC v. Catholic Univ. of Am., 83 F.3d 455, 463 (D.C. Cir. 1996) (employee was minister where her “primary functions serve [the religious employer’s] spiritual and pastoral mission”); Dayner v. Archdiocese of
What about claims sounding in contract or tort? Hosanna-Tabor expressly did not address such claims, not least because those claims were not presented in the case. But courts before 2012 had long held that the ministerial exception applies to a variety of claims, including condition-of-employment claims (such as wage-and-hour claims and hostile work environment claims), contract disputes, sexual harassments suits, and a number of tort claims such as tortious interference with business relationships, intentional infliction of emotional distress, invasion of privacy, and defamation. As the First Circuit explained thirty years ago, the most important consideration is not a claim's basis in contract, tort, or nondiscrimination law, but rather its "substance and effect" on the church's freedom to select and control its ministers. The court recognized that, in theory, some contract claims might be barred. If the answer is yes, then the claim will be more likely to be barred.

D. How Does the Ministerial Exception Operate at a Practical Level?

The ministerial exception's rule against state entanglement in internal church affairs is further safeguarded in a number of procedural ways. These safeguards provide a type of buffer around typical state powers employed during litigation, such as discovery requests and subpoenas, to prevent them from creating church-state conflict by being employed to probe the mind of the church. Recognizing that the "very process of inquiry" can "impinge on rights guaranteed by the Religion Clauses," courts have long enforced safeguards rooted in the First Amendment to, for instance, forbid intrusive inquiries into "confidential

none of the leading post-Hosanna-Tabor cases have squarely addressed the application of the ministerial exception to a tort claim. But the first and only federal appellate court to consider the issue on the merits of a contract claim agreed with the First Circuit. In Lee v. Sixth Mount Zion Missionary Baptist Church, the Third Circuit found that the exception applied to bar a claim that a church lacked sufficient cause to terminate its senior pastor. The court recognized that, in theory, some contract claims might not implicate ecclesiastical matters or require interfering with the internal governance of the church. But, in practice, as the court explained, every court that had reached the merits had "applied the ministerial exception [to bar] a breach of contract claim alleging wrongful termination of a religious leader by a religious institution." Because the plaintiff's claim required second-guessing the basis for the church's decision to terminate him, the Third Circuit found that it was barred under the ministerial exception's rule against entanglement.

Professor Douglas Laycock, who successfully argued Hosanna-Tabor and filed an amicus brief in support of the church in Sixth Mount Zion, appears to basically agree with the Third Circuit's approach. He believes that a "contract claim for unpaid salary or retirement benefits" can "surely" survive the ministerial exception. But a minister's breach of contract claim that disputes adequacy of cause is "squarely within the rationale of Hosanna-Tabor" and must be rejected:

A minister discharged for cause, suing in contract on the theory that the church lacked adequate cause to discharge him . . . would be directly challenging the church's right to evaluate . . . its own ministers, and he would be asking the court to substitute its evaluation of his job performance for the church's evaluation.

Professor Laycock's view, then, essentially tracks pre-Hosanna-Tabor case law: whether sounding in tort, contract, nondiscrimination law, or otherwise, the fundamental issue is whether a claim requires governmental entanglement with a church's sincere religious judgment about its relationship with its minister. If the answer is yes, then the claim will be more likely to be barred.

100  903 F.3d at 113. See Sixth Mount Zion, 903 F.3d at 120-122 nn. 5 & 7.

101  Id.

102  903 F.3d at 122.

103  The court explained that this non-entanglement principle was derived from the Establishment Clause component of the ministerial exception. But its ruling also suggested that, while previous Third Circuit precedent had indicated that a religious employer can contractually waive its Free Exercise rights under the ministerial exception, it was not clear that this ruling survived Hosanna-Tabor. See Sixth Mount Zion, 903 F.3d at 120-122 nn. 5 & 7.

104  Laycock, supra note 94, at 861.

105  Id.

communications among church officials.” 107 It was thus “well established” at the time of *Hosanna-Tabor* that state power should be sparingly employed to “troll[ ] through a person’s or institution’s religious beliefs.” 108 So how do those protections play out after *Hosanna-Tabor*?

### 1. Discovery

Post-*Hosanna-Tabor*, courts have consistently found that the ministerial exception is a threshold question that should be resolved before allowing discovery into the merits of claims that would necessarily fail should the defense succeed. Courts explain that this is crucial because unnecessary merits discovery creates “the very type of intrusion that the ministerial exception seeks to avoid,” 109 thus “making the discovery . . . process itself a first amendment violation.” 110 For those same reasons, courts have quashed subpoenas where they implicated or violated church autonomy rights. 111

Notably, this does not mean that courts deny all discovery. While plaintiffs can plead themselves out of a case, 112 courts will allow discovery where there are factual disputes related to the viability of the ministerial exception defense, with other merits discovery delayed until after that threshold issue is resolved. 113 Further, courts may also allow discovery into other claims that fall outside the reach of the ministerial exception. 114 In these respects, courts are largely following the general procedural approach for discovery related to jurisdictional challenges under Federal Rule of Civil Procedure 12(b)(1).

#### 2. Waiver

Courts have also faced several variations on the question of whether a party can waive the ministerial exception. The first case to squarely face the question following *Hosanna-Tabor* was the Sixth Circuit’s *Conlon* decision. There, the plaintiff argued that the defendant ministry had waived the ministerial exception by posting an equal opportunity employment policy expressly stating that the ministry would not discriminate on any bases other than religion. Thus, the plaintiff reasoned, the ministry waived any ministerial exception defense to her sex and marital status discrimination claims. The Sixth Circuit acknowledged that its precedent from 2007, *Hollins v. Methodist Healthcare, Inc.*, had indicated that waiver was possible where it was sufficiently express and unequivocal. But the Sixth Circuit found that *Hosanna-Tabor* abrogated *Hollins*, replacing it with a rule that the ministerial exception operated as “a structural limitation imposed on the government by the Religion Clauses, a limitation that can never be waived.” 115 The court explained that *Hosanna-Tabor* did “not allow for a situation in which a church could explicitly waive this protection” because the “protection is not only a personal one; it is a structural one that categorically prohibits federal and state governments from becoming involved in religious leadership disputes.” In short, a church cannot waive the government’s interest in remaining separate from the church. *Conlon* noted that the Seventh Circuit had reached the same conclusion in *Tomic v. Catholic Diocese of Peoria*, ruling that even if the parties should invite government involvement, a “federal court [should] not allow itself to get dragged into religious controversy.” 116 Three years later, in 2018, the Seventh Circuit’s *Grussgott* decision reaffirmed *Tomic*, finding that a religious employer’s equal opportunity policy cannot waive the ministerial exception. 117

Waiver came up in a different way in the Third Circuit’s *Sixth Mount Zion* decision. There, the district court raised the ministerial exception sua sponte after the defendant church failed to raise it in response to the senior pastor’s wrongful termination claim. On appeal, the Third Circuit affirmed the district court’s decision to do so, citing *Conlon* to support the conclusion that the church had not waived the ministerial exception “because the exception is rooted in constitutional limits on judicial authority.” 118

In another 2018 case, *EEOC v. R.G. & G.R. Harris Funeral Homes*, the Sixth Circuit again reaffirmed that the

---

107 Universidad Cent. de Bayamon v. NLRB, 793 F.2d 383, 401-02 (1st Cir. 1985) (en banc) (Breyer, J., concurring).


109 Sterlinski v. Catholic Bishop of Chi., 2017 WL 1550186, at *5 (N.D. Ill. May 1, 2017). See also Fratello, 863 F.3d at 190 (noting that the district court had restricted discovery).

110 Dayner, 23 A.3d at 1200 (citing McClure, 460 F.2d at 560 (“investigation and review” of the church’s relationship with its ministers would “cause the State to intrude upon matters of church administration and government which have so many times before been proclaimed to be matters of a singular ecclesiastical concern”)).


113 Herzog v. St. Peter Lutheran Church, 884 F. Supp. 2d 668, 671 (N.D. Ill. 2012) (allowing “limited discovery to determine whether the ministerial exception applies”).

114 See Dkt. 35, Garrick v. Moody Bible Institute, No. 18-cv-573 (N.D. Ill. July 3, 2018) (granting motion to reconsider order permitting discovery into merits of claims that would be foreclosed if ministerial exception defense was upheld, but permitting discovery into a claim that the court believed would survive the defense).

115 Conlon, 777 F.3d at 836.

116 442 F.3d 1036, 1042 (7th Cir. 2006).

117 Grussgott, 882 F.3d at 658.

118 Sixth Mount Zion, 903 F.3d at 118 n.4 & 121 (the doctrine is a “structural” limitation imposed on the government that safeguards courts from being “impermissibly entangle[d] . . . in religious governance and doctrine”). See also Bush v. Nation of Islam, 248 F. App’x 331, 333 (3d Cir. 2007) (affirming district court’s sua sponte raising of defense). But see Hamilton v. Southland Christian Sch., Inc., 680 F.3d 1316, 1319 (11th Cir. 2012) (finding that defendant had waived the ministerial exception defense by failing to raise it on appeal, but also noting that “[n]ot only did [defendant] fail to argue in its brief that the ministerial
ministerial exception operates as a structural limitation on court involvement in religious matters, regardless of whether either party raises it. There, as in Sixth Mount Zion, neither party argued that the ministerial exception was applicable to an employment discrimination claim before the court. But amici did raise it. Treating the exception almost as a jurisdictional issue which a court has independent responsibility to consider, the Sixth Circuit fully evaluated whether the ministerial exception applied and thus would prevent it from reaching the merits of the parties’ claims and defenses.

3. Interlocutory Appeal

Prior to Hosanna-Tabor, state appellate courts regularly permitted ministerial exception arguments to be raised on interlocutory appeal. The courts repeatedly emphasized that part of the ministerial exception right is protection against unnecessary litigation over and discovery into internal church affairs, and that interlocutory appeal was therefore necessary to vindicate that part of the right. For instance, the Connecticut Supreme Court warned that “the very act of litigating a dispute that is subject to the ministerial exception would result in the entanglement of the civil justice system with matters of religious policy, making the discovery and pre-trial process itself a first amendment violation.” By way of explanation, courts regularly compared church autonomy defenses to qualified immunity, a threshold legal issue that must be decided as a matter of law at the outset of a case and subject to appellate review when denied.

In the wake of Hosanna-Tabor, courts have continued to reach the same result. For instance, the Kentucky Supreme Court has repeatedly permitted interlocutory appeals of ministerial exception defenses. In 2014’s Kirby v. Lexington Theological Seminary, the court explained that “the determination of whether an employee of a religious institution is a ministerial employee is a question of law . . . to be handled as a threshold matter,” and interlocutory appeal is required both to ensure that the defense is “resolved expeditiously at the beginning of litigation” and to remove “the possibility of constitutional injury” that could otherwise be caused by discovery and trial.

Academics writing on the meaning and application of Hosanna-Tabor agree that denial of a ministerial exception defense “is effectively final and should ordinarily be permitted to be tested on interlocutory appeal.” They also agree that a ministerial exception defense “closely resembles qualified immunity” for purposes of the doctrine that permits such immunity claims to receive interlocutory appeal.

Since Hosanna-Tabor, federal appellate courts haven’t yet directly addressed the question in the context of a ministerial exception interlocutory appeal. But they have relied on Hosanna-Tabor and its church autonomy principles to permit related interlocutory appeals. The Seventh Circuit’s 2013 decision in McCarthy v. Fuller accepted an interlocutory appeal of a district court ruling that required the jury to decide whether the plaintiff was a member of a Roman Catholic religious order. The court explained that the First Amendment’s rule against judicial interference in internal religious affairs was “closely akin” to a type of “official immunity,” since it conferred “immunity from the travails of a trial and not just from an adverse judgment.” The court further explained that the “harm of such a governmental intrusion into religious affairs would be irreparable, just as in the other types of case[s] in which the collateral order doctrine allows interlocutory appeals.”

And in 2018, the Fifth Circuit granted interlocutory appeal of a district court order requiring Texas’ Catholic bishops to turn over internal church communications to abortion providers. There, the court held that it had jurisdiction to hear the appeal because “the consequence of forced discovery” on rights that “go to the heart of the constitutional protection of religious belief and practice” would be “effectively unreviewable” without an interlocutory appeal. The court relied on Hosanna-Tabor to note that “religious organizations” had an interest in “maintain[ing] their internal organizational autonomy [] from ordinary

---

119 884 F.3d 560.
120 Id. at 581-583.
122 Dayner, 23 A.3d at 1199-1200; White, 571 A.2d at 792-93 (“The First Amendment’s Establishment Clause and Free Exercise Clause grant churches an immunity from civil discovery and trial under certain circumstances.”).
123 Dayner, 23 A.3d at 1198-1200; Heard v. Johnson, 810 A.2d 871, 876-77 (D.C. 2002) (the ministerial exception is a “claim of immunity from suit under the First Amendment” that is “effectively lost if a case is erroneously permitted to go to trial”). See also Petrucco, 462 F.3d at 302-03 (making comparison to qualified immunity); Bryce v. Episcopal Church in the Diocese of Colo., 289 F.3d 648, 654 (10th Cir. 2002) (same).
124 426 S.W.3d at 604, 608-09. See also Edwards, 2018 WL 4628449, at *3 (permitting interlocutory appeal of ministerial exception defense).
125 Mark E. Chopko, Still a Threshold Question: Refining the Ministerial Exception Post-Hosanna-Tabor, 10 FIRST AMEND. L. REV. 233, 294 (2012) (denial of the ministerial exception defense “is effectively final and should ordinarily be permitted to be tested on interlocutory appeal”).
127 714 F.3d 971, 974-76 (7th Cir. 2013) (Posner, J).
128 Id.
129 Id. at 976.
130 Whole Woman’s Health, 896 F.3d at 368.
131 Id. at 367-68.
discovery.”132 The abortion providers sought en banc review and then certiorari to overturn the court’s exercise of interlocutory jurisdiction, but to no avail.133

IV. Conclusion

Courts can’t second-guess churches’ judgments of who should be their ministers. As Judge Robert Sack noted for the Second Circuit, courts are “[a]rmed only with the law as written and the tools of judicial reasoning,” leaving them “ill-equipped” to gainsay that, “for instance, a stammering Moses was chosen to lead the people, and a scrawny David to slay a giant.”134 Hosanna-Tabor’s unanimous vindication of the founders’ dual protection for internal religious autonomy and church-state non-entanglement avoids this problem. Courts are generally taking the Supreme Court’s cue, robustly applying Hosanna-Tabor’s reasoning to refine the law’s definitions of “ministry” and “minister,” as well as to prevent interference with church affairs through contract, tort, and procedural means.

132 Id. at 374.


134 Fratello, 863 F.3d at 203.
More News on Powers Reserved Exclusively to the States

By Robert G. Natelson

Federalism & Separation of Powers Practice Group

About the Author:
Professor Natelson taught constitutional law at the University of Montana and is Senior Fellow in Constitutional Jurisprudence at the Independence Institute in Denver. He has written extensively on constitutional law issues, particularly the Constitution's original meaning. For a list of publications, see https://i2i.org/constitution/articles-books-by-rob-natelson/.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:

This essay updates and supplements an article published last year in the Federalist Society Review entitled The Founders Interpret the Constitution: The Division of Federal and State Powers.1 That article explained how during the Constitution's ratification debates (1787-90), leading Federalists (the Constitution's advocates) issued authoritative enumerations of powers that would remain outside the federal sphere under the Constitution if ratified. Most of the enumerators were highly respected American lawyers. The two most important non-lawyers were Tench Coxe and James Madison. Coxe was a Philadelphia businessman and economist, member of the 1789 Confederation Congress, and future assistant secretary of the treasury.2 Coxe's ratification-era writings were highly influential among the general ratifying public—perhaps as influential as the essays in The Federalist.3

Subsequent interpreters of legal texts generally give considerable weight to representations of meaning presented by a measure's sponsors.4 The Federalists enumerating powers the Constitution denied to the central government clearly intended that the ratifying public rely on their representations. These representations squarely contradict claims by some commentators that the Constitution conferred near-plenary authority on the federal government.

This essay serves two purposes. First, it briefly addresses and refutes claims that near-plenary federal power lurks within two seemingly straightforward constitutional grants: the Commerce

---


3 Professor Cooke observed that, "Although Coxe's essays were not in the same literary league [as The Federalist], they perhaps were contemporaneously more influential, precisely because they were less scholarly and thus easier for most readers to follow." Cooke, supra note 2, at 111.

4 In founding era interpretation, as today, representations of meaning by a measure's sponsors carried far more weight than allegations by opponents. Such representations bound the sponsors later. Relevant legal maxims were Nemo contra factum suum venire potest ("No one may benefit [literally, "come"] in violation of his own deed"), Nemo potest mutare consilium suum in alius injuriisam ("No one may change his plan [or "advice"] to the injury of another), and Nullus commodum capere potest de injuria sua propria ("No one may benefit from his own injury").

Maxims of construction enjoyed great deference during the founding era. I Thomas Wood, An Institute of the Laws of England 6 (10th ed. 1772) ("[Maxims] are of the same Strength as Acts of Parliament when once the Judges have determined what is a Maxim"). An early American court accepted this view in State v.—, 2 N.C. 28, 1 Hayw. 29 (1794) ("And maxims being foundations of the common law, when they are once declared by the Judges, are held equal in point of authority and force to acts of Parliament").

---
Clause and the Necessary and Proper Clause. Second, it summarizes how materials reproduced in three newly published volumes in the Documentary History of the Ratification of the Constitution of the United States reinforce the conclusion of last year’s article.5

I. THE CONSTITUTION DID NOT GRANT NEAR-PLENARY AUTHORITY TO THE FEDERAL GOVERNMENT

The Constitution is notable for its grants of power. A natural reading of those grants seems to offer little justification for many of the activities of the modern federal government. For example, it is hard to see how a power to impose taxes6 includes authority to operate the Medicare program or shape the nation’s system of public education. But apologists for an expanded federal role have long offered broad interpretations of the Constitution’s grants to justify that role. Perhaps the first to do so was Alexander Hamilton, who in 1791 argued that the congressional power to tax to “provide for the general Welfare”7 authorized spending of any kind Congress thought served the general welfare.8 Notably, however, Hamilton refrained from offering this theory to the public until after the Constitution had been safely ratified; indeed, during the constitutional debates he argued to the contrary.9

The political environment during and after the New Deal encouraged expansive reinterpretations. Academics and judges defended the federal government’s newly broadened scope and sought ways to support it constitutionally. During that period the Supreme Court adopted Hamilton’s post-ratification reading of the General Welfare Clause.10 Commentators also began to argue for a more expansive definition of the Constitution’s phrase “to regulate Commerce.”11 Some contended the phrase encompassed not merely “commerce” in its strict sense (i.e., mercantile trade), but the entire national economy.12 The New Deal Supreme Court eventually adopted a variation of that view, although based more on the Necessary and Proper Clause than the Commerce Clause.13

More recent commentators have argued that the founding-era term “Commerce” was not limited to economic activities but referred to intercourse of all kinds, and that the clause therefore authorizes federal activity beyond what a natural reading would indicate. By this expanded reading, the Commerce Clause presumably authorizes Congress to regulate even what eighteenth century speakers sometimes called “commerce between the sexes.” That contention has not been made directly, but it has been claimed that the Commerce Clause empowers Congress to regulate interstate externalities of all kinds.14

Still other commentators have attributed a very broad scope to the Necessary and Proper Clause.15 For example, one writer argues that the provision in the clause reading “all other Powers vested by this Constitution in the Government of the United States” refers to a capacious and undefined inherent sovereign authority created by the Declaration of Independence and passed through the Continental and Confederation Congresses to the “Government of the United States.”16

II. EXPANSIONARY CLAIMS FOR THE COMMERCE AND NECESSARY AND PROPER CLAUSES ARE IMPLAUSIBLE

Readers familiar with the ratification record may notice a historical irony: The interpretative claims made by proponents of “big government” are eerily akin to those made by the Antifederalists, with their frenzied fears that the Constitution would result in a federal government out of control.17 During


13 Wickard v. Filburn, 317 U.S. 111 (1942); United States v. Darby Lumber Co., 312 U.S. 100 (1941). Wickard is by far the more famous decision, probably because of its memorable facts, but the conclusion in Wickard was dictated by the conclusion in Darby. The Court renders its reliance on the Necessary and Proper Clause, as opposed to the Commerce Clause, more explicit in Gonzales v. Raich, 545 U.S. 1 (2005).


15 U.S. CONST. art. I, § 8, cl. 18 (“The Congress shall have Power . . . To make all Laws which shall be necessary and proper for carrying into Execution the foregoing Powers.”).


17 For example, Professor Mikhail’s characterization of the Necessary and Proper Clause as a “sweeping clause,” Mikhail, Sweeping, supra note 16, cribbs from the Antifederalist playbook. E.g. 3 THE DEBATES IN THE SEVERAL STATE CONVENTIONS ON THE ADOPTION OF THE FEDERAL CONSTITUTION 150, 423, 436 et passim (Jonathan Elliot ed., 1866) [hereinafter Elliot’s Debates] (reporting this characterization by Patrick Henry).

Antifederalists called the Necessary and Proper Clause a sweeping clause to persuade the public that it granted powers beyond those enumerated, and that it therefore should not be ratified. However, it was not really a sweeping...
the ratification debates, the Antifederalists’ interpretations were refuted easily—at least as an intellectual matter—partly because they so often relied on rhetorical abuses such as wrenching constitutional phrases out of context and inserting words absent from the text. Moreover, educated readers could see that the interpretations Antifederalists offered bore little resemblance to how people actually wrote or read legal documents.

A. Interpreting the Commerce Clause

Consider, for example, the constitutional terms “commerce” and “to regulate commerce.” The meaning of those terms to the founding generation has been examined in three comprehensive studies published since 2001, together relying on several thousand eighteenth century usages. These studies found that, while broader meanings of “commerce” did exist, the word nearly always referred to mercantile trade and certain accepted incidents. They also show that “regulating commerce” was an established and discrete division of the law—like many other words and phrases defining the scope of the Constitution’s power grants: “Bankruptcy,” “Naturalization,” “establish Post Offices,” “Offenses against the Law of Nations,” and others.

Specifically, “to regulate commerce” meant primarily to set the rules for the body of law known as the law merchant. This was the jurisprudence governing mercantile trade and certain recognized incidents, such as commercial paper and marine insurance. In addition, the term “regulate commerce” included regulation of navigation and, to a lesser extent, other means of commercial carriage. However, the body of law labeled “regulation of commerce” certainly did not include governance over other activities affecting commerce or affected by commerce. This is why the framers enumerated separately congressional powers over such subjects as bankruptcy and intellectual property.

B. Interpreting the Necessary and Proper Clause

By an honest reading, the role of the Necessary and Proper Clause also was intended to be modest. Its wording was fairly typical of provisions in eighteenth century instruments granting enumerated powers. Provisions that were drafted as the Necessary and Proper Clause was drafted granted no authority at all; they were recitals explaining that the powers specifically listed carried incidental authority. This meant that a person granted enumerated powers could execute those powers by undertaking either (1) subordinate activities by which the enumerated powers customarily were executed or (2) subordinate activities without which execution of the enumerated powers would be very difficult.

These historical facts have not prevented at least one commentator from claiming the clause was far more ambitious than that. He argues that the phrase “the Government of the United States” tells us the government enjoys extraconstitutional inherent sovereign authority. During the Confederation Era, James Wilson claimed Congress enjoyed this kind of authority because he was frustrated with the strictly limited grants in the Articles of Confederation. The commentator contends that such authority passed to the newly formed federal government.

Of course, the Constitution derived its legal force from the ratification, so we must ask whether the ratifiers accepted that view. For many reasons, the answer is “no.” Wilson’s theory of inherent sovereign authority was widely loathed. His earlier clause as founding-era law used the term. A sweeping clause (also called a “sweeping residuary clause”) conveyed items additional to those enumerated (although of the same general kind) to prevent accidental omission. Moore v. McGrath [K.B. 1774] 1 Cowp. 10, 12, 98 Eng. Rep. 939, 941 (Lord Mansfield); cf. Strong v. Teatt [K.B. 1760] 2 Burr. 910, 922, 97 Eng. Rep. 628, 634. By its wording (reinforced by Federalist representation), the Necessary and Proper Clause adds nothing to the powers the Constitution otherwise grants. Infra note 28 and accompanying text.

18  E.g., Bruto, V. N.Y.J., Dec. 13, 1787, 14 DOCUMENTARY HISTORY, supra note 5, at 432, 433 (rewriting the General Welfare Clause to grant Congress “an authority to make all laws which they shall judge necessary . . . to promote the general welfare” and not mentioning that the General Welfare Clause pertained only to taxes); Centinel, V. Phila. INDEPENDENT GAZETTEER, Dec. 4, 1787, in 14 DOCUMENTARY HISTORY, supra note 5, at 343 (rewriting the Necessary and Proper Clause to authorize any law “Congress may deem necessary and proper”) (italics added).


20 U.S. CONST. art. I, § 8, cl. 4.

21 Id., art. I, § 8, cl. 4.

22 Id., art. I, § 8, cl. 7.

23 Id., art. I, § 8, cl. 10.

24 Natelson, Commerce, supra note 12. There is no need to be deterred by Professor Balkin’s statement that “the trade theory [of commerce] remains ad hoc and formalistic.” Balkin, supra note 14, at 22. The interpretive methodology modern law professors deride as “formalism” simply refers to the methodology dominant before they invented “legal realism.” The Constitution is primarily a formalistic document, intended to be construed in a formalistic way.

25 Natelson, Commerce, supra note 12, at 843.

26 U.S. CONST. art. I, § 8, cl. 4 (“The Congress shall have Power . . . To establish . . . uniform Laws on the subject of Bankruptcies throughout the United States”).

27 Id., art. I, § 8, cl. 8 (“The Congress shall have Power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”).


29 Mikhail, Sweeping, supra note 16.

advocacy of it damaged his popularity and rendered him an object of suspicion during the ratification debates. Accordingly, when those debates took place, Wilson and other Federalists took considerable pains to assure the ratifying public that the federal government would not have extensive implied authority. As part of these reassurances Wilson, like other Federalists, specifically and repeatedly denied that the Necessary and Proper Clause conveyed additional powers. So thoroughly did Americans reject the theory of inherent sovereign authority that they adopted the Tenth Amendment partly from a (vain!) hope that the theory would never plague constitutional discourse again.35

Did, nevertheless, Wilson and a few nationalist allies accomplish their true object by adding the “Government of the United States” phase to the Necessary and Proper Clause?36

Again, no. Even if Wilson and a few other framers secretly had that goal, it would be irrelevant. We don’t construe a document according to a secret intent not disclosed to, or shared by, those who were parties to the document. What is determinative is not what Wilson privately thought or hoped, but what the ratifiers were told and understood.37

Even those few founders friendly to the general notion of inherent sovereign authority would not have found the concept in the Necessary and Proper Clause. As noted above, the claimed textual hook in the clause is the reference to powers vested in “the Government of the United States.” The argument is that (1) because the Constitution’s other provisions do not grant power to the U.S. government as an entity, (2) to comply with the constructional preference against surplus, (3) we should assume the powers thereby referenced in the Necessary and Proper Clause derive from outside the Constitution.38 However, the first premise is wrong. The Constitution does contain provisions granting powers to the government as an entity. There is no need to posit an extra-constitutional source.

Several clauses grant such power, although they do so with language of obligation rather than language of grant. If Jill’s boss tells her, “You must represent our company in negotiating the Smith contact,” imposition of this mandate—words of obligation such as “must”—carries with it power to discharge it. Jill’s boss need not add, “I give you power to negotiate the Smith contract,” because the grant of power is implicit in the delivery of the mandate. This was true during the founding era as well.39 For example, the Crown’s instructions to colonial governors—which, along with the accompanying commissions, were precursors of the Constitution’s Article II—granted extensive powers to their recipients almost entirely through language of obligation.41 Indeed, the “take Care” formula common in colonial instructions reappears in the Constitution,42 where it grants the president authority to enforce the law even without reference to the Executive Vesting Clause.43 Similarly, Article I requires each house of Congress to keep and publish a journal of its proceedings,44 without any other language empowering each

31 Antifederalists dubbed Wilson, who was born and raised a Scot, “James de Caledonia,” and accused him of designs to create an all-powerful central government. E.g., James De Caledonia to James Bowdoin, Independent Gazetteer, Mar. 4, 1788, reprinted in 34 Documentary History, supra note 5, at 969.

32 E.g, James Wilson, State House Yard Speech, Oct. 6, 1787, https://www.constitution.org/asf/wilson0.htm:

‘[T]he congressional power is to be collected, not from tacit implication, but from the positive grant expressed in the instrument of the union. Hence, it is evident, that in the former case everything which is not reserved is given; but in the latter the reverse of the proposition prevails, and everything which is not given is reserved.

33 2 Elliot’s Debates, supra note 17, at 468 (James Wilson, speaking at the Pennsylvania ratifying convention) (“even the concluding clause, with which so much fault has been found, gives no more or other powers; nor does it, in any degree, go beyond the Particular enumeration”). See also id. at 448.


36 Professor Mikhail lists instances in which he claims certain framers disguised an expansive meaning by clever omissions and other devices. See Mikhail, The Necessary and Proper Clauses, supra note 16, at 1130-31.


38 Mikhail, Sweeping, supra note 16, at 40 (arguing that because the Constitution does not contain grants to “the Government of the United States,” “we must assume it refers to implied powers inherited from the Constitution’s congressional predecessors”). See also Mikhail, Necessary and Proper, supra note 16, at 1047.

39 Founding era drafters did not need to add separate power granting language to the language of obligation under at least three rules: (1) Necessarium est quod non potest aliter se habere (roughly, “If something in existence couldn’t exist without a thing, then that thing necessarily exists”), (2) Cuinqueque alquiis quid concedit concedere videtur et id sine quo res ipsa esse non potuit (“To whoever a person grants something is also granted that without which the grant cannot exist”), and (3) Frustra sit per plura, quod fieri potest per pacium (“It useless to establish by more words what can be established by many fewer.”). On the importance of the rules of construction during the founding era, see Natelson, supra note 30, at 28-31.


41 See, e.g., Instructions to Gabriel Johnson, available at https://i2i.org/wp-content/uploads/gabriel-johnston-instructions.pdf (Aug. 2, 1733) (relaying heavily on words of obligation as vehicles for empowerment of the royal governor of North Carolina); cf. Commission of Gabriel Johnson, available at https://i2i.org/wp-content/uploads/gabriel-johnston-draft-commission.pdf (May 10, 1733) (providing a more scanty list of powers). The form of the commissions was highly standardized. See Anthony Stokes, A VIEW OF THE CONSTITUTION OF THE BRITISH COLONIES 149-64 (1783). Both the commissions and the instructions were recognized as sources of authority. Id. at 183-84. See also id. at 199 & 202 (stating that the power to probate wills derives from a governor’s instructions).

42 U.S. Const. art. II, § 3 (listing among other presidential powers and obligation the duty to “take Care that the Laws be faithfully executed”).

43 Id., art. II, § 1.

44 Id., art. I, § 5, cl. 5 (“Each House shall keep a Journal of its Proceedings, and from time to time publish the same”)

2019 The Federalist Society Review 95
house to do so. The authority is encompassed in the duty. The Constitution contains other examples as well of power granted in the form of a mandate. 45

Moreover, not all the Constitution’s obligations are imposed merely on named officers and agencies. Article IV levies three obligations on “the United States.” The obligations are to protect states from domestic violence, to protect them from invasions, and to guarantee them republican forms of government. 46 Article VI imposes yet another duty on “the United States”: to pay Confederation debts. 47 These mandates necessarily convey the powers necessary to execute them. 48

What does the Constitution mean by “the United States”? Although the Constitution occasionally uses that phrase to refer to the country as a whole, 49 more commonly it means the U.S. government, including all its officers and instrumentalities. For example, the original Constitution mentions “the Treasury of the United States,” meaning the U.S. government’s treasury. 50 Similarly, it refers to an “Officer under the United States,” 51 meaning U.S. government officers as opposed to state officers, and to the “Coin of the United States.” 52 The Seventh Amendment refers to “any Court of the United States,” 53 meaning a court that is an arm of the U.S. government, but not of a state government. The Tenth Amendment speaks of powers “not delegated to the United States,” 54 meaning to the government and its officers and instrumentalities. The meaning of “the United States” in Articles IV and VI also refers to the government and its officers and instrumentalities.

For the reasons outlined earlier, each of the obligations Articles IV and VI imposes on the U.S. government necessarily conveys to the government power to comply with that obligation. The premise behind the “implied sovereign authority” version of the Necessary and Proper Clause—that the Constitution does not convey powers to the government per se—is simply inaccurate. The efforts of modern commentators to find massive hidden reservoirs of federal authority lurking in the Constitution’s straightforward grants are no more persuasive than similar efforts by their Antifederalist predecessors.

III. Federalist Representations of Federal Limits

Theories of near-plenary federal power contradict numerous and repeated representations the Constitution’s advocates made to the ratifying public during the constitutional debates. I summarized these representations in The Founders Interpret the Constitution. 55 These representations were not merely statements of expectation. They were specific representations to the ratifying public that the items enumerated were outside the federal purview. To my knowledge, modern advocates of federal omnipotence have never acknowledged the existence of those representations, much less attempted to account for them.

IV. Contributions from the New Volumes of the Documentary History

Earlier this year, the Wisconsin Historical Society published three new volumes of the Documentary History of the Ratification of the Constitution. 56 These volumes contain documents published in Pennsylvania during the ratification era but not included in the Pennsylvania volume of the Documentary History issued in 1976.

As a substantive matter, the three new volumes offer no surprises. As far as expansive claims for the Commerce and Necessary and Proper Clauses are concerned, the new volumes merely contribute more disproving evidence. For example, the term “commerce” appears many times, and the definable usages are consistent with, or reinforce, a scope limited to mercantile trade. 57 A newly reproduced founding-era discussion of the Necessary and Proper Clause adds to the available documentation affirming the provision’s narrow purpose. 58

54 supra note 1.
55 DOCUMENTARY HISTORY, supra note 5.
56 Robert G. Natelson, New evidence on the “Power To . . . regulate . . . Commerce,” available at [https://丁23.org/new-evidence-on-the-power-to-regulate-commerce/](https://丁23.org/new-evidence-on-the-power-to-regulate-commerce/), (collecting examples). Among the many uses of “commerce” in these volumes, I have found only one where the meaning is arguably broader. In Foreign Spectator, Phila. Independent Gazetteer, Sept. 12, 1787, in 32 DOCUMENTARY HISTORY, supra note 5, at 157, 159, the author quotes another’s work in which “commerce” could be interpreted to include economic activities other than agriculture. This is not a necessary interpretation, however.
57 A Subscriber, Phila. Independent Gazetteer, October 19, 1787, reprinted in 32 DOCUMENTARY HISTORY at 422:

In the 8th section, the power of Congress is declared and defined in several particulars, but as it was impossible to make all the laws at one time, which might be necessary to provide for the modes of exercising those powers, there is a general clause introduced which is confined to the powers given expressly by this Constitution to the Congress. It is, “to make all laws which shall be necessary and proper for carrying into execution the foregoing powers and all other powers, vested by this constitution in the government of the United States, or in any department or office thereof.” This certainly is not so much power as every other legislative body on this continent has, for the powers of this Congress are confined to what is expressly delegated to them; and this clause for enforcing their powers is confined merely to such as are explicitly mentioned. Yet have the words been stretched and distorted by some writers so as to give a power of making laws in all cases whatever. Nothing betrays the base designs of a writer more than his perversion of a plain meaning, which he often does by laying hold of some words and dropping others so as to make the fairest conduct appear in a shape that itself abhors.

96 The Federalist Society Review Volume 20
The new volumes’ most significant contribution is reprinting four essays by Tench Coxe, all signed “A Pennsylvanian.”58 The editors of the Documentary History excluded these essays from the initial Pennsylvania volume (reproducing them only on unindexed microfilm), they remained unavailable to most people. Perhaps the editors excluded them because they were published after the Pennsylvania ratifying convention concluded. Whatever the reason, their exclusion was a shame. Coxe’s Pennsylvanian essays are among the most significant of all Federalist writings.

On December 18, 1787, nearly all the Pennsylvania ratifying convention delegates voting against the Constitution issued a public apologia entitled The Address and Reasons of Dissent of the Minority of the Convention of the State of Pennsylvania.59 The Dissent raised several arguments, but the core of its case—like the core of the Antifederalist case generally—was that the Constitution would enable the federal government to become too powerful. The Dissent cited the General Welfare and Necessary and Proper Clauses as potential avenues toward federal tyranny. It was extensively distributed, both in Pennsylvania and in other states.60

The month after the Dissent’s publication, pro-Constitution correspondents wrote to Tench Coxe urging a public rebuttal.61 Coxe responded with eight essays addressing the Dissent. The first was signed “Philanthropos,” the next three “A Freeman,” and the last four “A Pennsylvanian.”62 The four Pennsylvania articles appeared in the Pennsylvania Gazette on successive weeks: February 6, 13, 20, and 27, 1788. They could not, of course, affect the result in Pennsylvania, but they were disseminated throughout the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified.

Coxe responded with eight essays addressing the Dissent. The first was signed “Philanthropos,” the next three “A Freeman,” and the last four “A Pennsylvanian.”62 The four Pennsylvania articles appeared in the Pennsylvania Gazette on successive weeks: February 6, 13, 20, and 27, 1788. They could not, of course, affect the result in Pennsylvania, but they were disseminated throughout the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified. One vehicle was the Gazette itself, perhaps the most respected newspaper in the country, notably in states that had not yet ratified.

In anticipation of national circulation, Coxe had addressed his Pennsylvania articles “To the People of the United States.” These contributions were of good quality. Coxe’s leading biographer, Professor Jacob E. Cooke, described them as “Coxe’s most noteworthy contribution to the ratification debate,” adding that they “invite comparison to the best of the literature spawned by that controversy, including the Federalist essays . . . .”63 In his first two Pennsylvania articles, Coxe pointed out that all the ratifying convention Antifederalists were strong supporters of the controversial Pennsylvania state constitution. He contrasted this with the more bipartisan cast of the Federalists. He thereby sought to establish that Pennsylvania’s Antifederalists were narrow partisans clinging to a defective state charter. In the fourth essay, he addressed some of the opponents’ arguments about the structure of the new government.

For present purposes, the most important essay is the third. There, and to a certain extent in the fourth, Coxe rebutted the crux of the Antifederalist case: that the proposed Constitution granted the central government too much authority. Coxe itemized a great many functions the Constitution placed permanently outside the federal sphere. Of them he wrote, “The legislature of each state must possess, exclusively of Congress, many powers, which the latter can never exercise.” In the fourth essay, he emphasized that the central government would have no control over religion. But the third contains his principal list of powers reserved to the states. These included operations a government must undertake by reason of being a government (such as creating and abolishing state offices and constructing “state houses, town halls, court houses”). They also included most traditional police powers. The third essay went on to say, in relevant part, that:

The state governments can prescribe the various punishments that shall be inflicted for disorders, riots, assaults, larcenies, bigamy, arson, burglaries, murders, state treason, and many other offenses against their peace and dignity, which, being in no way subjected to the jurisdiction of the federal legislature, would go unpunished. They alone can promote the improvement of the country by general roads, canals, bridges, clearing rivers, erecting ferries, building state houses, town halls, court houses, market houses, county gaols [jails—ed.], poor houses, places of worship, state and county schools and hospitals. They alone are the conservators of the reputation of their respective states in foreign countries, by having the entire regulation of inspecting exports. They can create new state offices, and abolish old ones; regulate descents of lands, and the distribution of the other property of persons dying intestate; provide for calling out the militia, for any purpose within the state; prescribe the qualifications of electors of the state, and even of the federal representatives; make donations of lands; erect new state courts; incorporate societies for the purposes of religion, learning, policy or profit; erect counties, cities, towns and boroughs; divide an extensive territory into two governments; declare what offences shall be impeachable in the states, and the pains and penalties that shall be consequent on conviction; and elect the federal senators. These things and many more can always be done by the state legislatures. How then can it be said, that they will be absorbed by the Congress, who can interfere in few or none of those matters, though they are absolutely necessary to the preservation of society and the existence of both the federal and state governments.

In the executive department we may observe, the states alone can appoint the militia and civil officers, and commission the same. They alone can execute the state laws in civil or criminal matters, commence prosecutions,
order out the militia on any commotion within the state, collect state taxes, duties and excises, grant patents, receive the rents and other revenues within the state, pay or receive money from Congress, grant pardons, issue writs, licences & c. [etc.—ed.] among their own citizens; or, in short, execute any other matter which we have seen the state legislature can order or enact. In the judicial department every matter or thing, civil or criminal, great or small, must be heard and determined by the state officers, provided the parties contending and the matter in question be within the jurisdiction of the state. Hence our petit and grand juries, justices of the peace and quorum, judges of the common pleas, our board of property, our judges of oyer and terminer, of the supreme courts, of the courts of appeal, or chancery, will all exercise their several judicial powers, exclusive and independent of the control or interference of the federal government.66

By organizing Coxe’s text and rendering it into modern language, we arrive at the following list of powers reserved to the states by the Constitution:

• With minor exceptions, ordinary criminal law is an exclusive state responsibility. Reserved to the states is jurisdiction over “disorders, riots, assaults, larcenies, bigamy, arson, burglaries, murders, state treason, and many other offences against [the states’] peace and dignity, which, being in no way subjected to the jurisdiction of the federal legislature.” Only the states may “declare . . . the pains and penalties that shall be consequent on conviction.”

• The states control civil justice within state boundaries.

• Infrastructure is almost exclusively a state function: “general [i.e., non-post]67 roads, canals, bridges, clearing rivers, erecting ferries.”

• Education and religion are exclusive state responsibilities. Only states may establish “state and county schools” and “places of worship,” or “incorporate societies for purposes of religion, learning, policy or profit.”

• The states enjoy exclusive jurisdiction over their internal commerce and other businesses since only they may “erect[] market houses” issue licenses, and inspect exports.

• Social services and health care are reserved exclusively to the states, for only states may establish “poor houses” and “hospitals.”

• The states retain exclusive power over inheritance and over land within their own boundaries.

All of these functions—as well as other items Coxe listed elsewhere—would be exercised by the states “independent of the control or interference of the federal government.” This enumeration is entirely consistent with all others issued by the Federalists.68 How these representations—widely distributed and unquestionably relied on—can be reconciled with plenary interpretations of federal enumerated powers is impossible to say.

66 Post roads were intercity, limited access highways punctuated by stations called “stages” or “posts.” Interstate highways are their modern analogues. See Robert G. Natelson, Founding-Era Socialism: The Original Meaning of the Constitution’s Postal Clause, 7 Brit. J. Am. Legal Studies 1 (2018).

68 For example, the following enumeration appears in The Freeman I, Pa. Gazette, Jan. 23, 1788, reprinted in 15 Documentary History, supra note 5, at 453, 457-58:

1st. Congress, under all the powers of the proposed constitution, can neither train the militia, nor appoint the officers thereof.

2dly. They cannot fix the qualifications of electors of representatives, or of the electors of the electors of the President or Vice-President.

3dly. In case of a vacancy in the senate or the house of representatives, they cannot issue a writ for a new election, nor take any of the measures necessary to obtain one.

4thly. They cannot appoint a judge, constitute a court, or in any other way interfere in determining offences against the criminal law of the states, nor can they in any way interfere in the determinations of civil causes between citizens of the same state, which will be innumerable and highly important.

5thly. They cannot elect a President, a Vice-President, a Senator, or a federal representative, without all of which their own government must remain suspended, and universal Anarchy must ensue.

6thly. They cannot determine the place of chusing senators, because that would be derogatory to the sovereignty of the state legislatures, who are to elect them.

7thly. They cannot enact laws for the inspection of the produce of the country, a matter of the utmost importance to the commerce of the several states, and the honor of the whole.

8thly. They cannot appoint or commission any state officer, legislative, executive or judicial.

9thly. They cannot interfere with the opening of rivers and canals; the making or regulation of roads, except post roads; building bridges; erecting ferries; establishment of state seminaries of learning; libraries; literary, religious, trading or manufacturing societies; erecting or regulating the police of cities, towns or boroughs; creating new state offices; building public houses, public wharves, county gaols, markets, or other public buildings; making sale of state lands, and other state property; receiving or appropriating the incomes of state buildings and property; executing the state laws; altering the criminal law; nor can they do any other matter or thing appertaining to the internal affairs of any state, whether legislative, executive or judicial, civil or ecclesiastical.

10thly. They cannot interfere with, alter or amend the constitution of any state, which, it is admitted, now is, and, from time to time, will be more or less necessary in most of them.

For other lists, see Natelson, Founders, supra note 1; Natelson, Enumerated, supra note 1.
Whistling in *Chevron*land: Why Department of Labor Interpretations of the Sarbanes-Oxley Act Whistleblower Provisions Do Not Deserve Judicial Deference

*By Donn C. Meindertsma*

Abstract:

As a rule, a federal court defers to an agency’s reasonable resolution of ambiguities in a law administered by the agency. Courts nonetheless hesitate to defer if Congress did not intend the agency action in question to carry the force of law or when other markers signify that Congress did not grant authority to resolve ambiguities. Following the general deference rule, federal courts routinely defer to the Department of Labor’s interpretation of whistleblower protection laws the DOL administers. Courts unhesitatingly defer even when those laws permit whistleblower claim adjudication by either the DOL or a court, as does the Sarbanes-Oxley Act. Because Congress assigned co-adjudicative authority under the Act’s whistleblower provision to the DOL and the judiciary, and for pragmatic reasons, deference to the DOL in this context is inappropriate.

If *Chevron* walked into an ABA conference, everyone would know who it was; no name badge required, no need for let’s-get-acquainted small talk (“What’s your holding?”). Like Cher, *Chevron* circulates with mononymous renown, its reputation preceding it. Still, because this essay tackles *Chevron*’s application in a particular context, a short re-introduction is in order at the outset.

The question in *Chevron* was whether the Environmental Protection Agency (EPA) permissibly interpreted “stationary source” in the Clean Air Act. The term could mean either a solitary pollution-emitting apparatus (say, a smokestack) or a single-sited cluster of them (say, a factory). Which construction was “right” in the context of the EPAs pollution control programs was perhaps irresolvable.

The Court deferred to the meaning the agency gave the law because Congress had entrusted the administration of the Clean Air Act to the EPA, “stationary source” was capable of more than one meaning, and the EPA’s interpretation of the term was reasonable. The muscle in *Chevron* was its holding that courts are to assume that statutory ambiguity exposes a congressional intent that an administering agency may resolve the ambiguity, so long as the agency’s construction is reasonable.

*Chevron* deference creates a dichotomy in judicial approaches to statutory interpretation. On one hand, in cases involving administrative law, courts defer to reasonable agency interpretations of ambiguous laws. On the other hand, in cases involving statutory ambiguity, courts are to resolve it themselves.

Note from the Editor:

The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:


About the Author:

Donn C. Meindertsma is a partner in the Washington, DC office of Conner & Winters.
interpretations unless they conflict with the law’s plain text.7 On the other hand, in ordinary federal court disputes involving statutory claims, courts seek the correct or at least best meaning of ambiguous text.8 A court will employ tools of construction and perhaps consider signposts such as congressional intent and legislative history. The construction on which the court settles maintains force as “the law” unless reversed on appeal, overruled, or abrogated.

This binary arrangement seems straightforward. The courts or the agency—one or the other—has the institutional authority to say what the law is.9 Chevron has been described as “institutional law” in that it “assigns to the administration the conditionally authoritative task of interpreting ambiguous statutory law and accordingly orders courts to under-enforce it.”10

This essay addresses a snag in this binary approach illustrated by (but not exclusive to) Section 806 of the Sarbanes-Oxley Act,11 which is one of several federal whistleblower laws. Statutorily, § 806 prohibits covered employers from retaliating against employees for reporting corporate fraud. Procedurally, § 806 offers twin resolution paths. The complainant may choose to litigate his or her claim before an agency—the U.S. Department of Labor (DOL)—or in federal court. The fact that complainants may choose between the ‘Agency Track’ and the “Court Track” upsets our otherwise neat binary arrangement. The law is co-administered and two-headed; a Siamese statute, if you will.

This essay argues that courts do not owe deference to DOL constructions of this statute. Part I details the § 806 framework. Part II discusses the justifications for deferring to statutory interpretation by agencies and summarizes court decisions on deference in § 806 cases. Part III summarizes Supreme Court and appellate court decisions on deference in this context. Part IV argues that courts should not defer to DOL interpretations of § 806.12

---

7 E.g., United States v. City of Fulton, 475 U.S. 657, 666 (1986) (saying the court “must uphold” agency interpretation “if the statute yields up no definitive contrary legislative command” and the agency’s approach was reasonable); Rapanos v. United States, 547 U.S. 715, 739 (2006) (finding that an agency’s interpretation was not a permissible construction of the statute); Grand Trunk W. R.R. Co. v. DOL, 875 F.3d 821, 831 (6th Cir. 2017) (whistleblower case).

8 “The judicial task, every day, consists of finding the right answer, no matter how closely balanced the question may seem to be.” Antonin Scalia, Judicial Deference to Administrative Interpretations of Law, 1989 DUKE L.J. 511, 520 (emphasis in original).

9 David J. Barron & Elena Kagan, Chevron’s Nondelegation Doctrine, 2001 SUP. CT. REV. 201, 202 (Chevron represents an institutional choice “between agencies and courts in ultimately resolving statutory ambiguities.”).


11 18 U.S.C. § 1514A.


---

I. Section 806’s Dual Tracks

The complainant in each § 806 case must initially file any retaliation claim against the employer with the DOL.13 He or she may then choose to litigate the claim entirely within the agency. If so, following discovery and a hearing, an administrative law judge (ALJ) will apply the statute to the claim, resolving statutory ambiguities as may be necessary, and issue a recommended order. The ALJ’s decision is then subject to review by the DOL’s Administrative Review Board (ARB),14 which reviews questions of law de novo and issues the DOL’s final order.15 Lastly, either party may seek review of that order in the appropriate circuit court of appeals. At this stage, the Solicitor of Labor, defending the ARB’s order, will solemnly apprise the court that it must defer to the ARB’s construction of § 806.16 Chevron will be cited.17

Alternatively, the complainant may file his or her § 806 claim in federal district court after a waiting period; this refiling is known as “kicking out.”18 Kicking out the complaint terminates DOL involvement as the district court assumes the familiar role

---


16 E.g., Brief for the United States as Amicus Curiae, Lawson v. FMR, LLC, No. 12-3, p. 13 (Sup. Ct., filed Apr. 9, 2013) (“[T]he ARB’s resolution of any ambiguity in the phrase ‘an employee’ is ‘controlling’ as long as it is reasonable.”).

17 Id. Section 806 requires appellate courts to conform to the review provisions of the APA. 18 U.S.C. § 1514A(b)(2)(A) (incorporating 49 U.S.C. § 42121(b)(4)). The APA directs the judiciary to decide questions of law, a standard that may or not may be compatible with Chevron. See Perez v. Mortgage Bankers Ass’n, 135 S. Ct. 1199, 1211 (2015) (Scalia, J., concurring) (“‘Needless of the original design of the APA, we have developed an elaborate law of deference to agencies’ interpretations of statutes and regulations. Never mentioning [the APA’s] directive that the ‘reviewing court . . . interpret . . . statutory provisions,’ we have held that agencies may authoritatively resolve ambiguities in statutes.’”) (emphasis in original).

18 Section 806 complaints first go to the Occupational Safety and Health Administration (OSHA). OSHA will begin an investigation and (if efforts to settle the claim fail) will issue a preliminary determination. After that, either party may request a hearing. Wherever the DOL proceedings stand after the first 180 days—whether OSHA has completed its investigation or not—the complainant may move the case to federal court. The kickout provision, 18 U.S.C. § 1514A(b)(1)(B), states:

A person who alleges discharge or other discrimination by any person . . . may seek relief . . . by—

(A) filing a complaint with the Secretary of Labor; or

(B) if the Secretary has not issued a final decision within 180 days of the filing of the complaint and there is no showing that such delay is due to the bad faith of the claimant, bringing an action at law or equity for de novo review in the appropriate district court of the United States, which shall have jurisdiction over such an action without regard to the amount in controversy.
of applying a statute in a dispute between private parties. If a jury is demanded, the jury will act as factfinder; the court will instruct the jury on the law. Per § 806, the district court proceeding is de novo. If the DOL issued any findings between the initial filing of the complaint and the refiling in federal court, those findings become moot.19

On appeal of a district court’s judgment, the appellate court would ordinarily review questions of law de novo.20 But that is not what happens in § 806 cases. Rather than exercise their right and duty to declare what the law means,21 appellate courts consider how the ARB has construed § 806—not in the case before the court, since the complainant opted out of the agency proceedings, but in any prior ARB decision.22 The appellate court will apply the ARB’s construction, if reasonable, even if the district court reasonably interpreted the law otherwise.

II. Deference and Its Justifications

A. Chevron’s Kin

Chevron, of course, is “not the alpha and the omega of Supreme Court agency-deference jurisprudence.”23 The decision was not written on the proverbial blank slate.24 Earlier cases had produced a common law of deference,25 under which agencies were permitted to reasonably construe ambiguous statutory terms where Congress entrusted them to carry out federal programs.26 However, the deference framework that developed was “never that simple” and “subject to override by a mélange of factors, with no clear metric for determining how much or when those factors weigh in the balance.”27

Of note is Skidmore v. Swift & Co.28 The question there was whether the Fair Labor Standards Act (FLSA) required the employer to pay wages for waiting time. The Court observed that whether time is compensable is a question of fact29 and that Congress assigned this factfinding responsibility to the courts. Yet the Court also recognized that the FLSA established an agency Administrator who had considerable experience in ascertaining the compensability of waiting time. Accordingly, although the Administrator did not preside like a court over individual employer-employee wage disputes, his opinions were entitled to due consideration by the courts.30 Skidmore famously concluded that the degree of deference owed an agency “depend[s] upon the thoroughness evident in its consideration, the validity of its reasoning, its consistency with earlier and later pronouncements, and all those factors which give it power to persuade, if lacking power to control.”31

As some see it, Chevron’s unadorned formula, ambiguity (A) —> deference (D), freed courts from having to ascertain congressional intent on an agency-by-agency, statute-by-statute basis. But post-Chevron decisions altered the A —> D formula. Fifteen years on, United States v. Mead Corp.32 held that Chevron deference is due only if Congress gave the agency the authority to make rules carrying the force of law and its determination was an exercise of that authority.33 This added a prerequisite, an “x” factor, to the equation: x —> (A —> D).34 Otherwise, the agency’s interpretation is entitled to respect only to the extent it has the power to persuade.35

Later, National Cable & Telecommunications Ass’n v. Brand X Internet Servs. (Brand X) presented the question whether, if a

---

20 Chad M. Oldfather, Universal De Novo Review, 77 Geo. Wash. L. Rev. 308, 308 (2009) (De novo review of questions of law “has become an accepted truth, one of those things that every lawyer knows and has known for so long that we regard it as an unalterable feature of the legal landscape.”).
21 Here, the citation obligatory in any deference discussion to Marbury v. Madison, 5 U.S. (1 Cranch) 137, 177 (1803).
22 Infra notes 86-95.
23 Eskridge & Baer, supra note 4, at 1120.
25 “[J]udicial control over administrative action has been based principally on the common-law doctrine of the supremacy of law, the due process guaranty embodied in the Constitution, and court interpretations of the statutory authority of administrative agencies.” B. Putney, Judicial review of administrative action, CONG. Q. 1938 (Vol. II). See E.E. Altobellworth, Judicial Review of Administrative Action by the Federal Supreme Court, 35 Harv. L. Rev. 127 (1921).
26 E.g., SEC v. Chenery Corp., 318 U.S. 80, 94 (1943) (“If the action rests upon an administrative determination—an exercise of judgment in an area which Congress has entrusted to the agency—of course it must not be set aside because the reviewing court might have made a different determination. . . .”); Bates & Gulick Co. v. Payne, 194 U.S. 106, 109–110 (1904) (action of agency head “whether it involve questions of law or fact, will not be reviewed by the courts unless he has exceeded his authority or this Court should be of opinion that his action was clearly wrong”); cf. Decatur v. Paulding, 39 U.S. 497, 515 (1840) (“If a suit should come before this Court which involved the construction of any of these laws, the Court certainly would not be bound to adopt the construction given by the head of a department. And if they supposed his decision to be wrong, they would, of course, so pronounce their judgment.”).
court had already settled on the meaning of a statutory term, an agency in an unrelated case down the road could embrace a different interpretation. This might have seemed like a rhetorical question: once a court has spoken, how can a bureaucrat say the law means something else, unsettling precedent on which other courts and private parties may have relied? Yet Brand X held that Chevron deference was still owed to the agency's interpretation. “[A]llowing a judicial precedent to foreclose an agency from interpreting an ambiguous statute . . . would allow a court’s interpretation to override an agency’s. Chevron’s premise is that it is for agencies, not courts, to fill statutory gaps.”

The Court reasoned that deference should not depend on the happenstance of whether the court's or the agency's interpretation came first. One commentator points out that "This is a ‘WOW’ moment. Brand X is arguably the capstone of the Court's Chevron evolution: it works a wholesale transfer of statutory interpretation authority from federal courts to agencies."

Lastly, but significantly, the Court has made clear that deference is not owed if "there may be reason to hesitate before concluding that Congress has intended . . . an implicit delegation” from Congress to the agency to fill in statutory gaps. This exception has been applied where the issues are particularly important. Nonetheless, if Chevron's premise is that Congress implicitly intended an agency to resolve statutory ambiguities, indications refuting such an intent should always be considered—even in cases important only to the litigants.

Chevron and its kin allow an agency to change the previously decided-upon meaning of a statute—to alter the law when a reason for alteration it finds. In contrast, stare decisis and other principles normally preclude a court from doing so, even if everyone thinks a prior ruling has become obsolete. What to a court is durable and controlling precedent is, in the polished-terrazzo halls of a federal agency, something like putty.

B. The Justifications for Deference

An abundance of commentary addresses why courts owe (or don't owe) deference to agency constructions of ambiguous statutes. While an in-depth exploration of justifications is not needed here, a synopsis aids in understanding whether deference in § 806 cases is appropriate.

1. Agency Expertise

Not surprisingly, Chevron cited agency expertise as a justification for its holding. Technical issues predominated in the litigation. The oral argument, heavy on statutory minutiae, was tedious, if not tranquillizing. Three Justices recused themselves. One can easily imagine the shrunked contingent of the remaining Justices in post-argument conference conceding the limitations on their ability to rightly define “stationary source.”

---


36 545 U.S. 967 (2005). The FCC ruled that cable companies that sell broadband internet service do not provide telecommunications service as the Communications Act of 1934 defined the term. Previously, the Ninth Circuit had decided that cable modem service is a telecommunications service, and in light of this prior “binding” panel decision declined to uphold the FCC’s ruling. The Supreme Court reversed.

37 Id. at 982. However, an agency may not depart from a court’s interpretation if the court deemed the law unambiguous. Id. at 982-83. Cf. Gutierrez-Brizuela v. Lynch, 834 F.3d 1142, 1149 (10th Cir. 2016) (Gorsuch, J., concurring) (“[T]here’s an elephant in the room with us today. We have studiously attempted to work our way around it and even left it unremarked. But the fact is Chevron and Brand X permit executive bureaucracies to swallow huge amounts of core judicial and legislative power and concentrate federal power in a way that seems more than a little difficult to square with the Constitution of the framers’ design. Maybe the time has come to face the behemoth.”).

38 Mead, 545 U.S. at 983. In dissent, Justice Scalia called the Court’s willingness to let the executive reverse judicial rulings a “breath-taking novelty,” as well as “bizarre” and “probably unconstitutional.” Id. at 1017.

39 Abbe R. Gluck, What 30 Years of Chevron Teach Us About the Rest of Statutory Interpretation, 83 Fordham L. Rev. 607, 625 (2014). See also Gutierrez-Brizuela, 834 F.3d at 1143 (Gorsuch, J.) (“[J]udicial declarations of what the law is haven’t often been thought subject to revision by the executive, let alone by an executive endowed with delegated legislative authority.”) (emphasis in original).

40 FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 159 (2000) (citing Stephen Breyer, Judicial Review of Questions of Law and Policy, 38 Admin. L. Rev. 363, 370 (1986)) (“A court may also ask whether the legal question is an important one. Congress is more likely to have focused upon, and answered, major questions, while leaving interstitial matters to answer themselves in the course of the statute’s daily administration.”); Christensen, 529 U.S. at 596-97 (Breyer, J., dissenting).


42 Apologies to Shakespeare, Sonnet 116 (“Love is not love which alters when it alteration finds.”). While all law is fluid, agency-made law certainly is even less “an ever-fixed mark” or a “star to every wandering bark.”

43 Northern Ind. Pub. Serv. Co. v. Carbon County Coal Co., 799 F.2d 265, 274 (7th Cir. 1986) (“We do not believe that we have the power to declare a constitutional statute invalid merely because we, or for that matter everybody, think the statute has become obsolete.”).


45 Apart from these justifications are “legal reasons” for deference; the one given in Chevron was that Congress intended agencies to have the power to resolve ambiguities. In Justice Breyer’s view, that is what Chevron was all about: “Chevron made no relevant change. It simply focused upon an additional, separate legal reason for deferring to certain agency determinations, namely, that Congress had delegated to the agency the legal authority to make those determinations. Christensen, 529 U.S. at 596 (Breyer, J., dissenting) (emphasis added).

46 “[A] full understanding of the force of the statutory policy in the given situation has depended upon more than ordinary knowledge respecting the matters subjected to agency regulations.” Chevron, 367 U.S. at 844 (citing, inter alia, NBC v. United States, 319 U.S. 190 (1943)); id. at 865 (“[T]he regulatory scheme is technical and complex” and “[j]udges are not experts in the field.”).


Agency expertise is a time-honored and pragmatic justification for deference. An agency’s “power to persuade” the courts, in the verbiage of Skidmore, surely correlates with its subject-matter proficiency and the complexity of the issue in dispute. As a rule, the more technical the basis for an agency’s decision, the more likely courts will defer. Subject matter and real world expertise also favor deference, as does an agency’s familiarity with the history and purpose of the legislation. The Supreme Court has stated that “historical familiarity and policymaking expertise account in the first instance for the presumption that Congress delegates interpretive lawmaking power to the agency rather than the reviewing court.”

49 E.g., United States v. Moore, 95 U.S. 760, 763 (1877) (citations omitted) ("The construction given to a statute by those charged with the duty of executing it is always entitled to the most respectful consideration, and ought not to be overruled without cogent reasons. The officers concerned are usually able men, and masters of the subject.").

50 Skidmore, 323 U.S. at 139 (A DOL Administrator has "more specialized experience and broader investigations and information than is likely to come to a judge in a particular case.").

51 See Chevron, 467 U.S. at 848 ("The Clean Air Act Amendments of 1977 are a lengthy, detailed, technical, complex, and comprehensive response to a major social issue."). The Court similarly relies on agency expertise as a reason to defer to agency interpretations of regulations: “Agencies (unlike courts) have ‘unique expertise,’ often of a scientific or technical nature, relevant to applying a regulation ‘to complex or changing circumstances.’” Kisor v. Wilkie, 588 U.S. __, slip op. 17 (2019) (quoting Martin v. Occupational Safety & Health Review Comm’n, 499 U.S. 144, 151 (1991)).

52 "For the most part, when the Court perceives agency rulemaking as steeped in technical expertise . . . [it] continues readily to defer." Seth Waxman, The State of Chevron: 15 Years after Mead, 68 Admin L. Rev. Accord 1, 12 (2016). See also Am. Elec. Power Co. v. Connecticut, 564 U.S. 410, 428 (2011) ("The expert agency is surely better equipped to do the job than individual district judges issuing ad hoc, case-by-case injunctions. Federal judges lack the scientific, economic, and technological resources an agency can utilize in coping with issues of this order.") (citing Chevron).

53 United States v. Haggar Apparel Co., 526 U.S. 380, 394 (1999) ("The expertise of the Court of International Trade . . . guides it in making complex determinations in a specialized area of the law; it is well positioned to evaluate customs regulations and their operation in light of the statutory mandate to determine if the preconditions for Chevron deference are present."); Pension Ben. Guar. Corp. v. LTV Corp., 496 U.S. 633, 651 (1990) ("[T]he judgments about the way the real world works that have gone into the PBGC’s anti-follow-on policy are precisely the kind that agencies are better equipped to make than are courts. This practical agency expertise is one of the principal justifications behind Chevron deference.").

54 Moore, 95 U.S. at 763 (Administrators are “not unfrequently . . . the draftsman[s] of the laws they are afterwards called upon to interpret.”); Commodity Futures Trading Comm’n v. Schor, 478 U.S. 833, 845 (1986) ("An agency's expertise is superior to that of a court when a dispute centers on whether a particular regulation is 'reasonably necessary to effectuate any of the provisions or to accomplish any of the purposes' of the Act the agency is charged with enforcing; the agency's position, in such circumstances, is therefore due substantial deference."); Eskridge & L. Baez, supra note 4, at 1109; Sarah Zeleznikow, "Leaving the Fox in Charge of the Hen House": Of Agencies, Jurisdictional Determinations and the Separation of Powers, 71 NYU Annual Survey of Am. L. 275 (2016).

55 Martin, 499 U.S. at 153.

Brand X accorded deference in part because it permits an agency to save a statute from ossification by revising “unwise judicial constructions of ambiguous results.” Deference frees agencies to formulate, refine, and change policy unburdened by “static” judicial interpretations. This may be less a justification for deference than an axiom—i.e., agencies receive latitude so that they have the leeway to modify the law. In any event, the save-from-ossification reasoning is also about agency expertise. On a forward-looking basis, as agencies confront scientific or technological changes, or new legislative or economic developments, they can incorporate the new information in implementing the statute Congress assigned them to administer.

2. Separation of Powers

Separation of powers ideals and democratic principles also animated Chevron. Chevron observed that, while agencies are not directly accountable to the people, the executive is. Therefore, it is appropriate for the executive “to make policy choices and to resolve competing interests that Congress inadvertently or intentionally left to be resolved.” “[F]rom a separation-of-powers perspective, absent some strong indication to the contrary, questions of what a statute means and how it is best implemented are for the Executive, not the Judiciary.” Chevron deference thus imposes restraints on the judiciary that limit interference with the executive’s advancement of public policy.

3. Additional Justifications

Chevron did not promote additional justifications for deference, but others have. Some have argued that deference will promote uniformity among the courts as to the meaning of a law. Without deference, courts may reach multiple and perhaps conflicting views of the meaning of a law. With deference, on the other hand, the courts more likely will coalesce around the meaning chosen by the agency. This rationale was summarized in the government’s brief in Brand X. The brief urged the Supreme
Court to overturn the Ninth Circuit’s decision (to not defer to the agency) because it would “subject a single agency decision to differing standards of review, thereby producing unseemly races to the courthouse, unnecessary conflicts in the circuits, and unfortunate situations in which (absent this Court’s review) the meaning of federal statutes would be dispositively determined for the entire Nation by lone three-judge panels.”

Another justification is that deference might prod Congress to draft legislation more precisely. The few ambiguities a law contains, the less opportunity the executive branch will have to alter its meaning. There appears to be a lack of empirical evidence that deference improves Congress’s drafting skills. In any event, legislative drafters may prefer ambiguous terms in the hope of producing a bill bland enough to pass.

III. Deference and § 806

Why is the DOL involved in § 806 at all? Did Congress have a particular reason related to the prevention of securities fraud—the objective of the Sarbanes-Oxley Act—for assigning the DOL to handle claims of retaliation for reporting fraud? The short answer is no. The DOL lacks fluency in federal securities laws and regulations. The agency is not conversant in the types of fraudulent conduct (e.g., wire fraud) covered by § 806. Nothing about the substance of the Sarbanes-Oxley Act would have led Congress to hand § 806 cases to the DOL.

Rather, historically, when Congress included a discrete anti-retaliation provision within a larger regulatory program, it assigned the DOL to handle retaliation claims. For example, when drafting the Clean Air Act, Congress preferred a DOL forum for whistleblower claim resolution over the EPA. By the time Congress drafted the Sarbanes-Oxley Act in 2002, the DOL already had jurisdiction over many similar whistleblower protection provisions. While the type of whistleblowing § 806 protects is distinctive (i.e., reporting shareholder fraud), Congress followed its common practice of tasking the DOL to handle whistleblower claims. There is no evidence that Congress considered the DOL uniquely capable to handle securities fraud whistleblower retaliation claims.

A. Supreme Court Decisions

The Supreme Court has twice considered federal financial whistleblower laws, although deference did not feature prominently in either case. The question in Lawson v. FMR, LLC was one of statutory interpretation: whether § 806 narrowly protects only employees of publicly traded companies or more broadly extends to workers of private companies that contract with publicly traded companies. A Court majority favored the more expansive reading. Because the Lawson plaintiffs chose the Court Track, the DOL had not issued its own decision on whether they qualified as covered employees. Deference therefore was not an issue for the majority.

Not so with the dissent. The dissent found § 806 ambiguous but concluded that the DOL’s interpretation of the statute (in other cases) did not deserve Chevron deference. The dissent reasoned that the DOL’s authority to investigate and adjudicate § 806 claims did not justify deference because the Sarbanes-Oxley Act did not delegate to the DOL any authority to make rules carrying the force of law. Instead, the Act gave the SEC the power to make rules necessary to protect investors. “[I]f any agency has the authority to resolve ambiguities in § [806] with the force of law, it is the SEC, not the [DOL].”

The dissent approached the nub of the issue addressed in this essay:

That Congress did not intend for the Secretary [of Labor] to resolve ambiguities in the law is confirmed by § [806]’s mechanism for judicial review. The statute does not merely permit courts to review the Secretary’s final adjudicatory rulings under the Administrative Procedure Act’s deferential standard. It instead allows a claimant to bring an action in a federal district court, and allows district courts to adjudicate such actions de novo. The dissent concluded that “the muscular scheme of judicial review suggests that Congress would have wanted federal courts,
and not the Secretary of Labor,” to have the ultimate power to resolve ambiguities in § 806. So far, no court has picked up on the dissent’s argument.

In Digital Realty & Trust, Inc. v. Somers the Supreme Court resolved a circuit split on the meaning of “whistleblower” under the Dodd-Frank Act. The Act defines a whistleblower as an employee who has reported wrongdoing to the SEC, but some appellate courts had construed the law to cover employees who had not done so. Somers considered deference, albeit not to the DOL. Pursuant to its rulemaking authority, the SEC had issued a rule that the Dodd-Frank Act protects employees even if they did not report wrongdoing to the SEC. The Court concluded, however, that the law means what it says: an individual is only protected if he or she reported wrongdoing to the SEC. Because the Court found that the statute was “clear and conclusive,” not ambiguous, the Court did not defer to the SEC’s conflicting construction.

B. Appellate Court Decisions

Federal courts, as a general matter, unhesitatingly defer to DOL interpretations of § 806. When plaintiffs kick out their cases to federal court, they metaphorically bring along a crate containing all DOL precedent for the court to sift through and apply in the (supposedly de novo) proceedings.

Consider how courts accommodate the DOL’s vacillating characterization of protected “whistleblowing.” The ARB held in 2008 that § 806 requires an employee’s report to “definitely and specifically” identify wrongdoing—not any old gripe will do. That construction of the law became known as the Platone standard. After the 2008 elections brought a new administration (and new ARB members), the ARB changed course; its conclusion that § 806 does not require definitive and specific reports is known as the Sylvester standard. Appellate courts by and large deferred to the Platone standard while it was “the law,” and then accorded the same deference to the new Sylvester standard. This was true even if, in a Court Track case, the court was reviewing a federal district court’s judgment as opposed to (in an Agency Track case) a final order of the ARB.

Take Wiest v. Lynch. The complainant pursued the Court Track, but the district court dismissed his claim because his evidence did not meet the requirements of Platone. However, in unrelated litigation, the ARB had just recently embraced the new Sylvester standard. The Third Circuit held that the district court should have opened the metaphorical crate of ARB precedent and applied the new standard.

Nielsen v. AECOM Tech. Corp. considered the same issue but justified deference under Skidmore. The district court had dismissed a § 806 claim based on circuit precedent adopting Platone. The Second Circuit reversed because the DOL in the meantime had repudiated that standard. The court declined to address whether Chevron deference was due, in part because the dissenting opinion in the Supreme Court’s Lawson case questioned whether the DOL has interpretive authority under § 806. Nonetheless, the court found the DOL’s new Sylvester standard persuasive.

Rhineheimer v. U.S. Bancorp Investments, Inc. took yet another approach. As had the Second Circuit in Nielsen, the Sixth Circuit rejected its own earlier embrace of the Platone standard while adopting Sylvester as persuasive. The court went on to find the correct interpretation of § 806 based on the “text and design” of the law and the “well-established intent of Congress” for “a broad reading of the statute’s protections.” These cases illustrate courts’ readiness to defer to the DOL in § 806 whistleblower cases without considering the fact that the district court had de novo jurisdiction to adjudicate the claim.

---

78 Id. at 478. “[M]uscular scheme of judicial review” is perhaps not the best phrasing. Section 806 does not suggest a heightened standard of review; it authorizes federal district courts to adjudicate claims ab initio once the DOL waiting period ends.
82 Dodd-Frank whistleblower cases are adjudicated only in federal courts so the DOL was not involved in the dispute. 18 U.S.C. § 15 U.S.C. § 78u-6(b)(1)(B)(1).
84 Rule 21F-2; 17 C.F.R. § 240.21F-2(a)-(b).
85 Somers, 138 S. Ct. at 781-82.
86 E.g., Delnek, Inc. v. DOL, 649 Fed. App’x 320, 327-28 (4th Cir. 2016) (“We defer to the Board’s interpretation of § 1514A.”).
87 Platone v. DOL, 548 F.3d 322, 326 (4th Cir. 2008) (reasoning that “due deference” is to be accorded to the ARB’s interpretation of § 806); see aho Welch v. Chao, 536 F.3d 269, 276 (4th Cir. 2008) (same); German v. ARB, 265 Fed. App’x 317, 320 n.7 (5th Cir. 2008) (“It appears that Chevron deference is due, as the ARB is an adjudicative body, but we leave that question for another day”).
IV. Courts Do Not Owe Deference to DOL Interpretations of § 806

Deference is not due to DOL interpretations of § 806 or comparable whistleblower protection provisions. The justifications for deference do not apply in this context, and pragmatic considerations make deference inappropriate.

A. The Dual-Headed Supervision of § 806 Refutes Any Fiction that Congress Delegated DOL Lawmaking Power

1. Who “Administers” § 806?

_Chevron_ requires deference to an agency’s construction of a “statute which it administers.” Does the DOL administer § 806 in the _Chevron_ sense?

On the one hand, a case can be made that it does. Since the early 1970s, Congress has used whistleblower protections as a means to accomplish the objectives of expansive regulatory programs, particularly in the environmental protection arena. Congress began the tradition of assigning the DOL to handle these claims; courts became involved only on petitions for review of final DOL orders.

On the other hand, with its enactment in 2002, § 806 departed from tradition by including the kickout provision. Congress empowered federal courts to hear and resolve § 806 cases de novo. This overt alternative to agency adjudication undercuts a conclusion that the DOL is “the administrator” of § 806.

Assume, for example, that Employer lays off two employees at the same time, purportedly for the same reason. They each file factually similar and legally identical claims, the validity of which hinges on the meaning of a term in § 806. Employee A keeps the case in the DOL process; practically speaking, it will take years before the ARB issues a final order on the claim. Employee B opts for the Court Track, and Employer soon files a dispositive motion, the outcome of which hinges on the resolution of the ambiguous text.

If deference to the DOL were required, a reasonable course for the federal court would be to stay Employee B’s case until the ARB chooses its construction of the ambiguous term in Employee A’s case. After all, assuming it is reasonable, that construction must govern the courts—if, of course, deference is due. Yet the very reason Congress provided the kickout provision was to permit complainants to escape the laggardly DOL process for resolving § 806 claims and obtain speedier justice, so postponing federal court proceedings until the ARB gets around to interpreting the law would frustrate that goal. Instead, the federal district judge should review, interpret, and apply § 806 on a de novo basis. In short, it cannot be said that the DOL is “the administrator” of § 806 insofar as federal courts have equal authority to apply and interpret the law.

2. The Missing Mead “X Factor”

Section 806’s dual-headed structure also suggests that the statute lacks the _Mead_-required force-of-law “oomph.” Standing alone, the fact that the DOL adjudicates § 806 cases could indicate that Congress intended the DOL to issue determinations that carry the force of law. Moreover, the proceedings are relatively elaborate and formal. _Mead_ “recognized a very good indicator of delegation meriting _Chevron_ treatment in express congressional authorizations to engage in the process of . . . adjudication that produces . . . rulings for which deference is claimed.” The DOL gives “concrete meaning” to the provisions of § 806 “through a process of case-by-case adjudication.”

But so do federal courts. If relatively formal agency proceedings aid the DOL in giving concrete meaning to § 806, even more formal federal court proceedings serve that function. For this reason, the establishment of dual adjudication tracks in § 806 indicates that Congress did not intend for the DOL to have plenary, or even primary, authority to resolve statutory ambiguities and, in turn, did not intend courts to defer to DOL interpretations.

---

96 See supra note 12.

97 467 U.S. at 842-43.

98 See Epic Sys. Corp. v. Lewis, 158 S. Ct. 1612, 1629 (2018) (“[O]n no account might we agree that Congress implicitly delegated to an agency authority to address the meaning of a . . . statute it does not administer. One of _Chevron’s_ essential premises is simply missing here.”); Price v. Stevedoring Servs. of Am., 697 F.3d 820, 833 (9th Cir. 2012) (“Skidmore strongly suggests that it is an administrative entity’s statutorily delegated authority to administer a statute that qualifies it for any kind of deference in the first place.”) (emphasis in original).

99 E.g., 33 U.S.C. § 1367 (enacted as part of the Water Pollution Control Act of 1972).

100 Pub. L. 107–204, 116 Stat. 745 (2002); cf. _Lawson_, 571 U.S. at 436 (“Congress has assigned whistleblower protection largely to the [DOL].”)

101 Specifying de novo review may be an unambiguous command that federal courts defer to DOL constructions of the statute, or it may more narrowly require courts to discount any prior findings of fact. Cf. _Stone_ v. Instrumentation Lab. Co., 591 F.3d 239, 246-47 (4th Cir. 2009) (“[T]he statute expressly requires the district court to consider the merits anew . . . [D]eferring to the administrative agency, even if more efficient, is in direct conflict with the unambiguous language of the Sarbanes-Oxley Act.”).

102 See _Lawson_, 571 U.S. at 477 (Sotomayor, J., dissenting).

103 “[T]he Secretary believes that access to district courts under this provision is intended to provide the complainant with a speedy adjudication of his complaint.” Procedures for the Handling of Retaliation Complaints Under Section 806 of the Sarbanes-Oxley Act of 2002, 80 Fed. Reg. 11865, 11877 (Mar. 5, 2015).

104 _Mead_, 553 U.S. at 230 (“It is fair to assume generally that Congress contemplates administrative action with the effect of law when it provides for a relatively formal administrative procedure tending to foster the fairness and deliberation that should underlie a pronouncement of such force.”); _Nielson_, 762 F.3d at 219-20.

105 533 U.S. at 229.

106 INS v. Cardoza-Fonseca, 480 U.S. 421, 448 (1987). The Third Circuit in _West_ deferred to the DOL for this reason, 710 F.3d at 130-31 (quoting _Mead_, 553 U.S. at 229); see also Dietz v. Cypress Semiconductor Corp., 711 Fed. App’x 478, 482 (10th Cir. 2017) (“As for legal determinations, this Court affords administrative deference to the ARB’s statutory interpretations, as expressed in formal adjudications.”)

107 “[C]ourts should only provide such deference when the relevant power has been delegated by Congress (even if such delegation is only implicit). Correspondingly, such deference should be withheld when such
True, other federal laws assign responsibilities to agencies and courts. In the employment context, Skidmore provides a ready example. That case centered on the FLSA, which permits federal courts to resolve disputes between employees and employers; in those cases, judges say what the FLSA means, while the DOL Wage and Hour Administrator also interprets the law, not least in deciding whether the agency should seek to enjoin employers from violating the statute. Brand X also involved a statute capable of federal court and agency (FCC) interpretation, depending on the context of the claim.

However, the provisions at issue in Skidmore and Brand X did not grant an agency and the courts de novo authority to adjudicate the very same claims. Section 806 assigns exactly the same roles in resolving retaliation claims to the DOL and the courts. Unlike the FLSA, the statute does not divide authority; it grants coequal authority.

Finally, the DOL does not have substantive rulemaking authority under § 806 (or similar federal whistleblower laws), which further suggests that deference to the DOL is not warranted. The only DOL regulations relating to § 806 are procedural. As the dissent in Lawson noted, the Sarbanes-Oxley Act empowered the SEC, not the DOL, to promulgate substantive rules.

B. The DOL Lacks Deference-Worthy Expertise

As discussed, agency expertise generally favors deference. But the DOL does not have expertise that would warrant deference to its interpretations of § 806.

As an overarching point, the core issue in retaliation cases—whether the employee was punished for reporting wrongdoing—is not a technical or esoteric one. It is a question of fact. Applying the law to the facts is the DOL’s bread and butter in Agency Tracking whistleblower cases. The ARB’s final, factual determinations whether retaliation occurred are reviewed on appeal under the substantial evidence standard. The same thing happens in a federal court proceeding. It does not take an expert to resolve § 806 claims.

Nor does it take DOL expertise to tease out the meaning of ambiguous statutory text. For example, a complainant must establish that he suffered an adverse action. Section 806 makes it unlawful to “discharge, demote, suspend, threaten, harass, or in any other manner discriminate against an employee” for his protected conduct. These are not terms of art. They pop up throughout the federal code, including in laws that do not involve the DOL. If these commonplace terms seem hazy, a federal district judge is as able as the ARB to resolve ambiguities.

Technical expertise is also not a basis for deference to the DOL. True, technical issues may arise in determining whether a wrongdoing report constitutes a protected form of whistleblowing. Whistleblower provisions often are embedded in regulatory programs that have technical components. Among them are several environmental laws, as well as programs that regulate commercial atomic power, aviation, and rail transportation. The federal agencies with relevant expertise in these areas are, in turn, the EPA, the Nuclear Regulatory Commission, the Federal Aviation Administration, and federal transportation agencies. The DOL does not develop technical know-how in these fields.

Thus, while the Sarbanes-Oxley Act was designed to prevent shareholder fraud, and while § 806 protects employee reports of fraud, the DOL is not a storehouse of understanding on corporate fraud and has no insight in that field beyond that of a federal court. As the dissent in Lawson noted, corporate fraud is the SEC’s territory, not the DOL’s—just as aviation safety is the FAA’s domain even though the DOL administers the aviation

---

108 29 U.S.C. § 216(b) (providing private right of action).
109 The DOL has also issued regulations and interpretive guidance on wage and hour law, strengthening its claim to FLSA deference. Skidmore, 323 U.S. at 139–40.
111 The DOL has a function the courts do not: OSHA conducts a preliminary investigation when the complaint is filed. That difference does not alter the comparative responsibilities and capabilities of the courts to adjudicate claims. First, in some cases OSHA fails to complete an investigation before the complaint moves the case to federal court—in which case the DOL has not served an additional function. Second, even if OSHA issues a determination, it plays no role in the later ALJ hearing or federal court trial, either of which is de novo.
112 Procedures for the Handling of Discrimination Complaints, 69 FR 52104, 52104 (2004) (“The purpose of this rule is to provide procedures for the handling of Sarbanes-Oxley discrimination complaints; this rule is not intended to provide statutory interpretations.”).
113 571 U.S. at 477 (Sotomayor, J., dissenting).
114 At least to an extent, expertise is a prerequisite for deference to statutory interpretations. See Dantran, Inc. v. DOL, 246 F.3d 36, 48 (1st Cir. 2001) (“Agency regulations interpreting a statute that relates to matters outside the agency’s area of expertise are entitled to no special deference.”) (citing Adams Fruit Co. v. Barrett, 494 U.S. 638, 649–50 (1990)).
116 Id.
117 E.g., 42 U.S.C. § 7622 (Clean Air Act).
118 42 U.S.C. § 5851. See Kansas Gas & Elec. Co. v. Brock, 780 F.2d 1505, 1510 (10th Cir. 1985) (saying the court is “troubled” by DOL involvement in nuclear regulatory matters but noting that “substantial questions involving competence in nuclear energy are involved, the NRC may provide technical assistance” in whistleblower cases to the DOL).
122 Brown & Root, Inc. v. Donovan, 747 F.2d 1029, 1033 (5th Cir. 1984) (“Nuclear energy involves questions of great scientific and engineering sophistication well beyond that required in ordinary industrial relations. The Department of Energy (in particular, the Nuclear Regulatory Commission) has special competence in this area, not the Department of Labor.”).
123 Cf. Murray v. UBS Securities, LLC, 2013 WL 2190084 *7 (S.D.N.Y. 2013) (asserting that an SEC rule broadly interpreting § 806 reflects the “considerable experience and expertise that the agency has acquired over time with respect to the interpretation and enforcement of the securities laws”), abrog’d by Somers, 138 S. Ct. 767.
whistleblower protection provision.\textsuperscript{124} In any event, corporate fraud aptitude is not required to resolve § 806 claims.\textsuperscript{125}

Additional factors indicate that § 806 simply does not require agency expertise. First, the whistleblower laws assigned to the DOL are concise—usually contained in a single statutory subsection and covering no more than two or three pages. They are not intricate national programs requiring specialized agency knowhow.\textsuperscript{126} Second, whistleblower claims are a minor aspect of the DOL’s affairs. Sometimes, Congress establishes an agency for the very purpose of overseeing a federal program, as with the Social Security Administration.\textsuperscript{127} In that case, the overseeing agency will develop subject matter expertise. Congress did not establish the DOL to handle a complex, national, retaliation-prevention program.

The DOL’s familiarity with labor markets and statistics does not lend it mastery to resolve ambiguities in whistleblower laws.\textsuperscript{128} Section 806 is unlike the complex laws the DOL does not possess expertise to resolve ambiguities in whistleblower protection provision.\textsuperscript{124} In any event, corporate fraud aptitude is not required to resolve § 806 claims.\textsuperscript{125} It is true that bureaucratic pockets within the DOL develop a specialized level of “whistleblower law” comprehension. OSHA investigates § 806 claims and claims under about two dozen similar federal laws, under the supervision of a national director of whistleblower programs and regional whistleblower staff.\textsuperscript{132} DOL ALJs, too, may over time become proficient in applying whistleblower statutes in concrete cases, given that retaliation claims make up a sizeable portion of their dockets. Repetition nonetheless does not make ALJs relatively more competent than federal judges to resolve statutory ambiguities.\textsuperscript{133}

In fact, the DOL does not hold itself out as possessing an inherent capability to interpret whistleblower laws. ALJs and the ARB often look to court interpretations of terms commonly used in anti-discrimination and labor laws, such as Title VII of the Civil Rights Act\textsuperscript{134} and the National Labor Relations Act.\textsuperscript{135} Also, the DOL cannot claim to be better qualified than courts to construe general law terms, such as punitive damages and limitations provisions, simply because those terms appear in whistleblower laws.\textsuperscript{136}

\textsuperscript{124} “The DOL has been charged with administering whistleblower complaints in a variety of employment contexts, even where another agency, having the technical expertise in the subject area of the complaints (such as the SEC here), has overall control.” Carnero v. Boston Sci. Corp., 433 F.3d 1, 16 n.13 (1st Cir. 2006).

\textsuperscript{125} Day v. Staples, Inc., 555 F.3d 42, 55 (1st Cir. 2009) (“Fraud” itself has defined legal meanings and is not, in the context of § 806, a colloquial term.).


\textsuperscript{127} See Perez, 135 S. Ct. at 1222 (Thomas, J., concurring) (citing Marbury, 5 U.S. 137); Beck v. CNO Fin. Grp., Inc., 2018 WL 2984854 *4 (E.D. Pa. June 14, 2018) (nuclear whistleblower statute did not create a specific administrative body to handle claims but simply assigned DOL to do so, and DOL lacks any “special expertise” in resolving retaliation claims).

\textsuperscript{128} “The purpose of the [DOL] shall be to foster, promote, and develop the welfare of the wage earners of the United States, to improve their working conditions, and to advance their opportunities for profitable employment.” 29 U.S.C. § 551.

\textsuperscript{129} Cf. Long Island Care at Home, Ltd. v. Coke, 551 U.S. 158, 165 (2007) (FLSA case; “The subject matter of the regulation in question concerns a matter in respect to which the [DOL] is expert, and it concerns an interstitial matter, i.e., a portion of a broader definition, the details of which, as we said, Congress entrusted the agency to work out.”); Kasten v. Saint-Gobain Performance Plastics Corp., 563 U.S. 1, 14 (2011) (giving Skidmore deference to DOL in its interpretation of anti-retaliation provision of FLSA, 29 U.S.C. § 215(a)(3); id. at 23 (Scalia, J., dissenting) (deference inappropriate because DOL has “no general authority to issue regulations interpreting the Act, and no specific authority to issue regulations interpreting” the provision in issue)).

\textsuperscript{130} In contrast, the Supreme Court has recognized the special function of the National Labor Relations Board and Federal Labor Relations Authority to apply federal labor law to the complexities of industrial and federal labor relations. NLRB v. Erie Resistor Corp., 373 U.S. 221, 236 (1963); National Fed. of Fed’t Employees, Local 1309 v. Department of the Interior, 526 U.S. 86, 99 (1999).

\textsuperscript{131} Cf. Rhinehimer, 787 F.3d at 809-10 (noting in passing that, under Mead, agency expertise may warrant deference). Decisions under other federal whistleblower laws sometimes credit DOL expertise. They do so, however, in perfunctory manner, as if taking judicial notice of an inarguable truth, without exploring whether the claimed expertise is fact or fiction. E.g., United States v. Constr. Prod. Research, Inc., 73 F.3d 464, 472 (2d Cir. 1996) (nuclear whistleblower case; asserting that retaliation claims “are within the DOL’s particular area of expertise”); Bechtel Const. Co. v. Soc’y of Labor, 50 F.3d 926, 933 (11th Cir. 1995) (nuclear whistleblower case; “[T]he Secretary’s expertise in employee protection entitles his view to deference.”).

\textsuperscript{132} See supra note 18.

\textsuperscript{133} ALJs may, however, become adept with experience in making factual determinations. Pan Am Rys. v. DOL, 855 F.3d 29, 39-40 (1st Cir. 2017).

\textsuperscript{134} 42 U.S.C. §§ 2000e-2000e-17; Youngerman v. UPS, ARB No. 11-056, slip op. at 4 (ARB, Feb. 27, 2013) (“[W]e often look to Title VII precedent for guidance given the similarities in the anti-discrimination statutes.”). A court may refuse to enforce ARB determinations that depart from Title VII precedent. E.g., Stone & Webster Constr. v. DOL, 684 F.3d 1127, 1134-35 (11th Cir. 2012) (“The ARB failed to correctly identify and follow our circuit’s Title VII precedent…. [which] may not be binding, but the Secretary does not deny that her agency ‘routinely’ follows it.”).

\textsuperscript{135} 29 U.S.C. §§ 151–169.

\textsuperscript{136} Worcester v. Springfield Terminal Ry. Co., 827 F.3d 179, 182 (1st Cir. 2016) (according Skidmore deference to ARB’s application of punitive damages provision in rail safety whistleblower decision where the ARB had followed the reasoning of a Supreme Court case; see also City of Arlington, 133 S. Ct. at 1881 (Roberts, C.J., dissenting) (contending that deference is due only if Congress charged an agency to administer the specific statutory provision at issue).
“Some interpretive issues may fall more naturally into a judge’s bailiwick.”

Moreover, Congress has not given the DOL authority over all federal whistleblower laws. Federal courts have jurisdiction over whistleblower claims under the Dodd-Frank Act\(^\text{138}\) and False Claims Act.\(^\text{139}\) And the DOL has a limited role in adjudicating claims under the whistleblower provision of the Occupational Safety and Health Act, Section 11(c),\(^\text{140}\) which are by far the most common type of whistleblower claim that the DOL receives.\(^\text{141}\) Section 11(c) does not create a private right of action, and there are no administrative claims for the DOL to adjudicate.\(^\text{142}\) For all these reasons, the DOL cannot claim any § 806 expertise.

C. No Separation of Powers Concerns

Some judges, including those on the Chevron Court, and scholars posit that deference to the executive honors the Constitution’s separation of powers framework. One scholar argues that deference is a “soft constitutional norm” that encourages the judiciary to exercise restraint and to avoid dictating outcomes in policy-laden areas.\(^\text{143}\)

These significant, if lofty, ideals do not justify deference to the DOL in construing § 806. Congress explicitly gave the DOL and the courts the authority independently to adjudicate claims. Opting for the Court Track excludes the DOL from further considering a claim. Section 806 grants equal power to two branches, so courts have no reason to restrain themselves from deciding what the law means.

D. Pragmatic Concerns Counsel Against Deference to DOL Interpretations of § 806

Pragmatic justifications are relevant to the role of judicial deference. When an agency pursues policies based on “judgments about the way the real world works,” deference is owed for the very practical reason that the agency is “better equipped” to make such judgments.\(^\text{144}\) But as demonstrated above, the DOL cannot claim that it is better equipped than a federal court to interpret and apply § 806. In fact, pragmatic concerns counsel against deference.

1. Impracticality in Federal Litigation

Requiring a federal district court to scour the corpus of DOL caselaw before settling on the meaning of § 806’s terms seems a peculiar imposition. It is one thing to expect a court to weigh an agency’s interpretation of a statutory term on a petition for review, as is the case when a court of appeals reviews an ARB’s final order at the conclusion of an Agency Track case. In that setting, the reviewing court considers the ARB’s explicit interpretation of a particular term, with the benefit of the agency’s reasoning, in the fact-specific context of the case at hand.

In a de novo federal court proceeding, however, if DOL precedent governs, the court (and counsel) would need to master ARB precedent in case an ARB decision, at some point in the past, defined an ambiguous term relevant to the litigation. Jury instructions about the law might need to be rewritten each time the ARB resolves equivocal statutory language. A trial court’s failure to apply (or even notice) a statutory gloss the ARB adopted could be ground for reversal.\(^\text{145}\) Because federal courts are at least as equipped at the DOL to properly read § 806, they should not be regarded as lesser, secondary authorities on the meaning of the law.

2. Inconsistent Application of § 806

Skidmore accorded deference to the DOL in part because it believed agency and court interpretations of the FLSA should be uniform.\(^\text{146}\) Some posit that agency interpretations should prevail over a court’s (rather than vice versa) because deference will produce the happy result of court coalescence around the agency’s single interpretation. Without deference, different federal circuit courts might read § 806 in conflicting ways.

Flaws in the coalescence hypothesis are apparent. Deference will not promote uniformity when courts disagree about whether 1) a provision is ambiguous, 2) the agency’s determination has the “effect of law” per Mead, or 3) the agency’s interpretation is reasonable. The one way to assure uniformity—aside from Congress clarifying the law by amendment—is to percolate conflicts up to the Supreme Court. Lawson made the Court’s interpretation of § 806 the uniform “law of the land.” Deference had nothing to do with it.

Another barrier to coalescence is that the ARB has no commitment to stare decisis. Whatever courts decide, the ARB may change its mind. Political change in the executive branch leads to new, sometimes partisan ARB membership. Newly formed ARBs may be prone to quickly jettisoning “politically incorrect” decisions of the previous administration. Courts may coalesce around an interpretation, only then to coalesce around a different one. For example, perhaps the DOL’s current Sylvester interpretation of § 806 is correct. Maybe Platone better respects the statute. Or perhaps the ARB has not yet found the best interpretation of the law on this point. A new interpretation

---

\(^{137}\) **Kisor**, slip op. 17 (addressing deference to an agency’s interpretation of its regulations).


\(^{139}\) 31 U.S.C. § 3730(h). E.g., Halliburton, Inc. v. ARB, 771 F.3d 254, 267 (5th Cir. 2014) (finding § 806 language plain and essentially identical to statutory text in the False Claims Act).

\(^{140}\) 29 U.S.C. § 660(c).

\(^{141}\) See whistleblowers.gov/factsheets_page/statistics.

\(^{142}\) The DOL may bring Section 11(c) actions in federal court, in which case the courts will resolve statutory ambiguities, although some courts have deferred to the DOL’s views. In addition, in Whirlpool Corp. v. Marshall, 445 U.S. 1 (1980), the Court accorded Skidmore deference to a DOL regulation that interpreted Section 11(c).

\(^{143}\) Mark Seidenfeld, Chevron’s Foundation, 86 Notre Dame L. Rev. 273, 275 (2011).

\(^{144}\) **LTV Corp.**, 496 U.S. at 652.

\(^{145}\) Courts would need broad knowledge not only of the ARB’s interpretations of § 806 but also interpretations under analogous laws that track the terms of § 806. See supra note 12; **Lawson**, 571 U.S. at 431 (quoting S. Rep. No. 107-146) (Congress designed § 806 to track “as closely as possible” the aviation whistleblower law, 49 U.S.C. § 42121).

\(^{146}\) **Skidmore**, 323 U.S. at 140; Mead, 533 U.S. at 234 (noting the ”value of uniformity in [the] administrative and judicial understandings of what a national law requires”).
may surface as the DOL reshapes the living, breathing law.147 Coalescence, if any, will always be temporary and comes at the expense of finality.

3. Inconsistency with Other Employment Laws

In applying whistleblower statutes, the ARB may determine for any number of reasons to depart from the accepted meaning of terms routinely used in employment statutes. In United Turbines v. DOL, for example, the Second Circuit reviewed the ARB’s interpretation of the term “discharge.”148 That word is not a term of art and in fact appears in many federal statutes (rather unlike the term “stationary source”). The ARB decided that a “discharge” can include situations where the employer did not actually discharge the worker but erroneously believed the worker resigned. Postulating (without any illumination of the point) that the ARB “has a significant expertise in handling whistleblower claims,” the Second Circuit deferred to this outlier interpretation, even as the court observed that the ARB’s “reading does not mirror the definition that we have applied to similar terms in other employment laws.”149

In that instance, deference did not bring uniformity to the law. Deference by each of the other circuits to this odd interpretation might achieve uniformity within the narrow arena of § 806 cases (at least until the ARB changes its mind), but surely that is too modest a judicial goal. Uniformity in the broader arena of federal employment laws would benefit employers and employees alike. If the courts commonly believe discharge carries its usual meaning, allowing the DOL to part ways with the commonly accepted usage brings divergence.150 Except where the specific language in a statute calls for another interpretation, discharge should mean roughly the same thing for all employment laws.151

Finally, courts may foster uniformity by refusing to defer to the ARB. For example, the ARB once took it upon itself to apply a novel test for determining the liability of employers in harassment cases. This effort met a quick demise in the Fifth Circuit, which held that the DOL had no business—even in administering a law assigned to the agency—departing from the national understanding of workplace harassment liability.152

V. Conclusion

Perhaps statute-by-statute analysis of Chevron’s application is unwieldy. Yet the Supreme Court has made clear that there is more to the deference analysis than Chevron suggested. The Court sometimes finds reasons not to defer to agency interpretations of ambiguous statutory provisions, including where the matter at hand is simply too weighty to allow the agency to have the final word or where other factors counsel against deference. For the reasons outlined above, DOL interpretations of § 806 do not deserve deference. Federal courts should discontinue their habit of routinely affording deference and reclaim their authority to say what the law is.

147 The Supreme Court has required the DOL to explain new interpretations of the law where the change affects “serious reliance interests.” Encino Motorcars, LLC v. Navarro, 136 S. Ct. 2117, 2127 (2016).
148 581 Fed. App’x 16 (2d Cir. 2014).
149 Id. at 18. The court made no attempt to apply canons of construction to first determine that “discharge” is ambiguous.
150 The DOL also must “color within the common-law lines” when applying terms that have a common-law meaning, such as “employer.” Browning-Ferris Indus. of Calif. v. NLRB, 911 F.3d 1195, 1208 (D.C. Cir. 2018).
151 “It is a ‘fundamental canon of statutory construction’ that, ‘unless otherwise defined, words will be interpreted as taking their ordinary, contemporary, common meaning.’” Sandifer v. U.S. Steel Corp., 571 U.S. 220, 227 (2014) (quoting Perrin v. United States, 444 U.S. 37, 42 (1979)).
152 Williams v. ARB. 376 F.3d 471 (5th Cir. 2004); AKM LLC dba Volks Constructors v. Sec’y of Labor, 673 F.3d 752, 757 (D.C. Cir. 2012) (declining to accord deference to the DOL’s interpretation of a limitations provision in part because it “runs afield of our precedents”); cf. Worcester, 827 F.3d at 182 (accordign deference to DOL application of punitive damages standards because it conformed with broader Supreme Court precedent on punitive damages).
Have the American People Irrevocably Ceded Control of Their Government to the Modern Administrative State?

By Ted Hirt

Administrative Law & Regulation Practice Group

A Review of:

About the Author:
Ted Hirt is a Professorial Lecturer in Law at the George Washington University Law School. He previously was a career attorney in the United States Department of Justice's Civil Division for almost 37 years. The views he expresses are his own.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. We invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

The past few years have witnessed a surge of writing by conservative intellectuals about the modern administrative state, including how its expansive reach might be constrained or reversed.¹ The most recent contribution to this important development in our civic discourse is Professor John Marini's Unmasking the Administrative State: The Crisis of American Politics in the Twentieth-First Century. Marini contends that our modern centralized administrative state, with the active support of many prominent nineteenth and twentieth century social scientists, upset and supplanted America's original political theory of liberal constitutionalism, under which our nation had a limited government that distinguished between the public and private spheres and between the state and broader civil society.² Our nation's original theory of limited constitutional government, Marini argues, was based on "a reasonable and realistic understanding of the relationship of theory and practice, of ends and means."³ That understanding was based on the virtue of "prudence, not science," insofar as prudence "presupposes the possibility of moral virtue to direct men to the right, or good, ends."⁴ But the modern administrative state has replaced those principles with reliance on a technocratic bureaucracy that is convinced that rational administration can solve economic and social problems.⁵ This transformation has replaced the "sovereignty of the people" established in the Constitution with the "sovereignty of government," under the auspices of the modern rational administrative state.⁶

Marini is a professor of political science at the University of Nevada, Reno and a Senior Fellow of the Claremont Institute in California. He received his PhD in government at the Claremont Graduate University, and he also has taught at Ohio University and the University of Dallas. During the Reagan Administration, Marini served as a Special Assistant to then-Chairman of the Equal Employment Opportunity Commission, Supreme Court Justice Clarence Thomas.

The book's editor, Ken Masugi, a Senior Fellow at the Claremont Institute, says in his introduction that, in October 2016, Justice Thomas mentioned him and Marini as having

---


³ Marini, at 9.

⁴ Id.

⁵ Id. at 8-9.

⁶ Id. at 13.
been his first “mentors” on the Constitution. Unmasking the Administrative State is a compilation of Professor Marini’s essays and presentations, written or delivered during the past several decades. Although the collection contains material going back to the 1970s, it nevertheless has a contemporary focus. Two chapters offer Marini’s reflections on President Donald Trump’s successful 2016 election campaign and his presidency. Although the book’s essays address discrete political issues, Unmasking the Administrative State’s overall theme is that our modern administrative state is an unconstitutional centralization of political power in the federal bureaucracy.

I. How Did the Administrative State Triumph Over the Constitution?

Marini traces the establishment of the modern administrative state in the United States, in large part, back to President Franklin D. Roosevelt’s “reinterpretation” of the Constitution when his administration, supported by Congress, launched the New Deal programs. In a September 1932 speech, President Roosevelt asserted that the relationship of the government to the people was essentially contractual—“rulers were accorded power, and the people consented to that power on consideration that they be accorded certain rights.” That formulation enabled the government to determine the conditions of a new social compact, which diminished the authority of the Constitution and undermined popular sovereignty. This new understanding of the government as the “arbiter” of both economic and political rights enabled the government to place the expertise of the bureaucracy in charge of policymaking, thereby replacing the “moral authority of the people’s compact.”

Roosevelt’s political triumph had been preceded by decades of Progressive thinking that posited that rights were not natural or individual in origin, but instead were based in societal norms. The noted philosopher John Dewey criticized the founders for their belief that liberty is derived from natural rights, arguing that their understandings were “historically conditioned” and did not take into account the idea of “historic relativity.” In 1917, the eminent legal scholar Roscoe Pound observed that modern legal philosophy asked for “a definite, deliberate, juristic program as part of an intelligent social program, and expects that program to take account of the maximum of human demands and to strive to secure the maximum of human wants.”

American Progressivism, Marini contends, was the “political manifestation of a theoretical revolution in political thought,” derived ultimately from a “philosophy of History.” The German philosopher Immanuel Kant argued that the moral law could not be established on natural law or natural rights. Progressive intellectuals like Woodrow Wilson understood “natural laws only in terms of science, not ethics or morality,” and they concluded that the founders’ reliance on natural law principles was obsolete and had been superseded by scientific progress. The Progressives, influenced by German philosopher Georg Wilhelm Friedrich Hegel, also concluded that the modern state would become the vehicle for progress, with politics and religion replaced by the rational science of economics and society. American Progressive political scientists believed, like their European counterparts, in a theory of social justice under which the government would provide political solutions to contemporary social and economic problems.

Marini points out that one danger of the philosophy of History is that it obscures a correct understanding of the reality of tyranny” because the philosophy rests on the assumption that rationalism will triumph over time. The noted political scientist Leo Strauss observed that political science failed to recognize the persistence of tyranny across time, and that even though tyrannies like Hitlerism and Stalinism were destroyed, a modern tyranny that relies on science and technology remains an ongoing danger. A liberal or constitutional democracy that retains a limited government and the rule of law could be a bulwark against the growth of a modern, centralized administrative state that also could be tyrannical.

Marini contrasts the philosophy of scientific rationalism with the founders’ alternative vision of a moral law that is derived from the laws of nature. The principles of the Declaration of Independence and the political theory embodied in the Constitution rested on the idea that individual natural rights are the best basis to ensure the people’s sovereignty and security. In the Declaration, Thomas Jefferson invoked “the Laws of Nature and of Nature’s God,” and the Constitution’s Preamble

---

7 Id. at 1.
8 Id. at 2-3.
9 Id. at 29-39, 273-86.
10 Id. at 6-9, 13.
11 Id. at 15-17.
12 Id. at 16 (quoting from 1 Franklin D. Roosevelt, The Public Papers and Addresses of Franklin D. Roosevelt (1938)).
13 Id. at 17.
14 Id. at 17-18.
15 Id. at 18-20.
17 Id. at 22 (quoting Roscoe Pound, Juristic Problems of National Progress, 22.6 Am. J. of Sociology 721-33 (May 1917)).
18 Id. at 224.
19 Id. at 224-25.
20 Id. at 225-26.
21 Id. at 226-29.
22 Id. at 235.
23 Id. at 260-61.
24 Id. at 260-62.
25 Id. at 264-65.
26 Id. at 233.
27 Id. at 15.
emphasized that the people of the United States had “ordain[ed]” and “establish[ed]” the Constitution, making them sovereign.28 The founders believed that natural rights existed prior to the government’s creation, and that the government’s responsibility was to defend and secure those rights, not to create them.29 The founders intended that all three branches of the newly-formed government would derive their authority from the Constitution, and that the branches would exercise that authority on behalf of the common good.30

The Constitution therefore structured our federal government so that none of the branches could dominate the others, and so that political conflict would be regulated and resolved within that structure.31 There would be diverse views of the public good within the government—the separation of powers would ensure that the different branches would not coalesce around a single vision of the common good—and there would be “independent constituencies in support of each branch.”32 This would, by design, make it difficult for the government to create and justify a “unified will” as the basis for a political right to govern.33

In one essay, Tocqueville’s Centralized Administration and the “New Despotism,” Marini illustrates the tension between individual liberty and the pressure toward centralized government through the work of French writer Alexis de Tocqueville, who compared the early American government to the monarchies (and successor governments) of Europe.34 According to Marini, in The Old Regime and the Revolution and Democracy in America, Tocqueville observed a “democratic tendency” towards centralized government.35 The French Revolution accelerated that tendency and destroyed the elements of French society that could resist it.36 Tocqueville predicted—with prophetic insight—that socialism and centralization would “thrive on the same soil,” insofar as both ideologies advocate expansive government powers and the elimination of all class distinctions, resulting in a single authority over the public and no outlet for public opinion but the dominant state itself.37

Marini also observes that Tocqueville was impressed by early nineteenth-century America’s local decentralization, which facilitated a “civic spiritedness and love of liberty” that kept individuals from an exclusive focus on their self-interest.38 But Tocqueville was concerned that the loss of that localized power would enable a centralization that would lead inevitably to despotism.39 Tocqueville also observed that a nation’s laws need to preserve to each citizen a “political existence” that encompasses both rights and duties, with a resulting “civic conscience.”40 Functioning within local institutions, citizens can exert their efforts to maintain individual liberties while participating in government.41

Tocqueville identified an additional danger to democratic societies: the philosophy of human perfectibility, which, when combined with notions of equality, can result in a loss of individual identity and an obsession with unity.42 The democratic spirit, when combined with the idea that humans are “endowed” with an “indefinite faculty for improvement,” results in understanding equality as something to be implemented through “uniform human legislation.”43 Centralized government and an isolationist individualism are the ultimate results of such a philosophy.44 The two go together because isolation results in dependency on government rather than self-government.45 Equality “places men beside one another without a common bond to hold them.”46 Tocqueville urged a renewed commitment to individual liberty as the only antidote to a human tendency to accept centralization as a solution to civil society’s challenges.47 But Marini also senses in Tocqueville an almost fatalistic acquiescence to the process of centralization and bureaucratization, traceable, Marini contends, to Tocqueville’s acceptance of the view that “will had replaced reason as the distinctive characteristic of man.”48

Taken together, the founders’ principles and Tocqueville’s reflections on democracy and equality sharply differ from the Progressives’ “philosophy of History” under which the state would become the rational scientific planner of both law and politics.49 In a land governed according to the latter theory, the sovereignty of an enlightened people would be replaced by a government that was increasingly indifferent to their interests and no longer

59 Id. at 156.
60 Id. at 161 (quoting Roger Boesche, Tocqueville and Le Commerce: A Newspaper Expressing His Unusual Liberalism, J. OF THE HIST. OF IDEAS 44 (April-June 1983)).
61 Id. at 161-62.
62 Id. at 165-66.
63 Id. at 165-66 (quoting Alexis de Tocqueville, Democracy in America 427 (trans. Harvey Mansfield and Delba Winthrop 2000)) (hereinafter “Democracy in America”).
64 Id. at 172 (quoting Democracy in America, 641).
65 Id. at 173.
66 Id. (quoting Democracy in America, 485).
67 Id. at 174-75.
68 Id. at 184. Marini also discusses Tocqueville’s views in a related essay, On Harvey Mansfield’s Jefferson Lecture: How to Understand Politics, Id. at 177-84, in which Marini salutes the work of renowned political scientist Harvey Mansfield. Mansfield, like Tocqueville, links liberty to “human greatness.” Id. at 178.
69 Id.
needed to justify its legitimacy or mode of governance to them.\textsuperscript{50} Constitutionalism would thus become a historical anachronism.\textsuperscript{51}

Marini therefore poses the question of whether the modern administrative state has irrevocably undermined the principles of limited government, separation of powers, American federalism, and self-government.\textsuperscript{52} Marini traces our contemporary plight to our political institutions, and allied constituencies, which have adapted to the centralized bureaucracy and have permitted its continuous expansion.\textsuperscript{53}

II. Is Congress Complicit in the Expansion of the Administrative State?

Just as Progressive thinkers like Woodrow Wilson and political leaders like Franklin D. Roosevelt vigorously advocated creation of an administrative state to address perceived economic and social needs, Congress also has “enabled” the expansion of the administrative state.\textsuperscript{54} Congress initially resisted centralized administration because it wanted to keep its “deliberative, representative, and lawmaking functions.”\textsuperscript{55} With President Lyndon Johnson’s 1964 landslide election victory, however, Congress acquiesced to the expansion of social programs that destroyed federalism and undermined the separation of powers.\textsuperscript{56} With the enactment of Great Society legislation, Congress adapted to a new function as “guardian” of the administrative state.\textsuperscript{57} Congress and the president took for granted the legitimacy of the administrative state, and they no longer questioned whether its centralizing powers were consistent with the Constitution’s limits on such powers.\textsuperscript{58}

Marini is pessimistic as to the prospect that Congress will try to dismantle the administrative state.\textsuperscript{59} He believes that the Washington establishment has little incentive to change how it does business.\textsuperscript{60} One obstacle to Congress taking on this responsibility is the federal bureaucracy itself, which, Marini contends, has become an independent political faction.\textsuperscript{61} In addition, government is increasingly driven by the necessity of accommodating “various organized, political, economic, demographic, or social” groups that have “coalesced around the administrative state.”\textsuperscript{62} Furthermore, the bureaucracy can replace political decision-making “by substituting administrative rulemaking for general lawmaking, and rule by expert in place of that of elected official”;\textsuperscript{63} this enables legislators to avoid accountability to voters for their decisions, as they pass the buck along to unknown administrators who do not have to face reelection.\textsuperscript{64} The representative role of political parties in the national legislative process also has diminished; “bureaucratic patronage became more important than party patronage.”\textsuperscript{65} Members of Congress—and the interest groups that interact with them—have determined that it is more efficient to effect (or resist) policy changes by advocating their views directly to agency officials rather than making their case to the American people.\textsuperscript{66} This focus has contributed to a centralization of policymaking.\textsuperscript{67}

In addition, members of Congress face an inherent conflict between advancing the interests of Congress as a functioning, legislating body, and advancing their self-interest by serving the parochial interests of their districts.\textsuperscript{68} The challenge of reinvigorating Congress may thus depend in part on how individual members can be incentivized to make their success more dependent on “institutional performance and less dependent on their personal efforts.”\textsuperscript{69}

Congress has also ceded to the federal judiciary the sole responsibility to determine the legitimacy of administrative actions.\textsuperscript{70} The judiciary, in turn, has deferred to Congress’s decisions to delegate wide swaths of its authority to the agencies.\textsuperscript{71} The courts, by not holding Congress responsible for enacting “purposely unfinished laws,” allow agencies and affected interest groups to negotiate the rules that govern our society.\textsuperscript{72} The satisfaction of interests replaces the rule of law.\textsuperscript{73} The political branches and the national political parties, “organized around the private interests of national elites,” have thus created a centralized administrative state that includes various kinds of elites but excludes the broader electorate. The broader electorate, in turn, can access the government only through the political parties, but the parties no longer serve as a true “link” between the people and their government.\textsuperscript{74}

In his 1959 book, Congress and the American Tradition, James Burnham expressed concern about Congress's ability to
survive in an era of executive “dominance.” Burnham advocated a strong Congress that would craft public policy rather than allow the executive branch and administrative agencies to make those judgments. The problem, Marini observes, is that while Congress has retained its autonomy and authority, it has engaged in a “wholesale delegation” of power to administrative agencies. Congress is therefore the agencies’ “overseer,” with committees and individual members primarily engaged in overseeing the departments and agencies. And, although Congress has strengthened its oversight of agencies, individual members only intervene in the “execution phase” of the governing process, leaving policymaking still in agency hands.

Marini identifies several events that contributed to the dilution of congressional authority. For example, after the 1994 midterm elections, then-Speaker Newt Gingrich and his supporters altered the committee system to centralize control from his office, weakening “deliberation, representation, and the accommodation of interests that culminate in lawmaking on behalf of a public good.” With reduced membership participation and centralization of authority in the Senate and House majority leadership staffs, there is less expertise, more influence by private stakeholders, and ultimately a strengthened administrative state.

Marini also situates historic controversies in the context of America’s age of centralism. President Richard Nixon clashed with Congress, ultimately leading to his resignation in 1974. Most Americans probably associate this clash with the 1972-74 Watergate scandal, which culminated in the House of Representatives voting to impeach President Nixon for complicity in covering up the June 1972 Watergate Hotel burglary. For Marini, however, the clash between President Nixon and Congress must be understood in a different context—the legitimacy of presidential power in national politics. President Nixon claimed that his 1972 election victory was a mandate to curb the federal bureaucracy and centralized power more generally. Marini argues that the bureaucracy itself played a substantial role in the Watergate crisis.

In his 1972 reelection campaign, President Nixon deployed the increasing growth of the size and power of the centralized administration. Nixon’s solution was to further centralize executive power into the White House and away from the “permanent government.” Nixon also intended to reverse the flow of power to Washington by restoring decision-making to the states and localities. Nixon felt a personal mandate, arising out of his landslide victory over Senator George McGovern, to exercise power as president to achieve those difficult objectives—objectives that certainly would be opposed by political opponents, organized interest groups, elites, and the national media.

Nixon promptly articulated his goal of reversing the “the age of centralism” in American government. In his January 5, 1973 message to Congress, Nixon deployed the “balkanization of the departments and agencies,” and the loss of independence of state and local governments. Nixon earlier had attempted to reorganize the executive branch by abolishing several agencies and consolidating their functions. Nixon’s efforts in 1971
encountered resistance, so, in 1973, he turned to his own executive powers to accomplish the reorganization of executive agencies.\textsuperscript{99} Nixon’s budget for fiscal year 1974 also tried to restore more authority to states and localities.\textsuperscript{100} Congress, however, did not react favorably—it counterattacked.\textsuperscript{101} In February 1974, the Senate voted to require Senate confirmation of the Budget Director, a position that had been filled by presidential appointment, without Senate confirmation, since its creation 52 years before.\textsuperscript{102} Thus, one legacy of the Watergate scandal was that a presidential effort to assert control over the administrative state failed.\textsuperscript{103}

Marini’s most provocative essay addresses the influence of the ideology of the administrative state on our nation’s immigration policies.\textsuperscript{104} At first (or second) glance, it may be difficult for the reader to understand the relevance of immigration policy to the administrative state. But Marini succeeds in showing how this contentious issue fits within his overall narrative.

Marini contends that the problem of immigration is not intelligible unless one understands what constitutes “the ground of unity or common identity” as a nation.\textsuperscript{105} In the founding era, the United States was identified as a “regime of civil and religious liberty,” unlike European nations where a common religion was the original basis of citizenship.\textsuperscript{106} In 1790, President Washington wrote his famous letter to the Hebrew Congregation in Newport, Rhode Island, in which he identified the United States as the nation in which everyone could “possess alike liberty of conscience and immunities of citizenship,” and in which all could exercise “their inherent natural rights,” and the government simply required that citizens “who live under its protection should demean themselves as good citizens in giving it on all occasions their effectual support.”\textsuperscript{107} President Abraham Lincoln later reflected that, although nineteenth-century immigrants were not related to the founders by blood, the principles of the Declaration of Independence united immigrants with the native born.\textsuperscript{108} America’s social compact was founded on ideas of freedom and opportunity rather than birth and privilege.\textsuperscript{109}

The founders’ ideas of natural rights and the social compact that provided the early American understanding of citizenship and immigration were fundamentally altered by the acceptance of Progressive-era thinking.\textsuperscript{110} Progressive intellectuals like Herbert Croly viewed the ideal democracy as one in which an individual would “serve the nation in the very act of contributing to his own individual fulfillment.”\textsuperscript{111} John Dewey viewed individuals and society as “organic to each other,” with the state representing that organized relationship.\textsuperscript{112} These thinkers and others rejected the idea of a social compact made by individuals, thus repudiating Lincoln’s idea of a Union built on that compact.\textsuperscript{113} Ironically, Marini notes, some post-Civil War intellectuals, in rejecting Lincoln’s equality principle, also endorsed theories of race and color that would restrict immigration to northern Europeans.\textsuperscript{114}

Marini thus characterizes the restrictive Immigration Act of 1924, which imposed strict national and group quotas on immigration, as the culmination of the almost fifty years of an ideology that “celebrated the rational state as the embodiment of the moral will of a people,” which will had now been defined by “blood, race, class, or culture.”\textsuperscript{115} In 1916, the \textit{New Republic} editorialized that freedom of migration from one country to another country was an element of nineteenth-century liberalism “that is fated to disappear.”\textsuperscript{116}

Immigration policy shifted again, however, with the Immigration and Nationality Act of 1965.\textsuperscript{117} The national origins quota system was replaced by individual criteria, with an emphasis on admitting immigrants based on their skills.\textsuperscript{118} But Marini argues that both Presidents John F. Kennedy and Lyndon B. Johnson would have understood that the new immigrants would be “shaped by the expectations created by government, and not those of a free society,” thus making those immigrants constituencies for the Democratic Party.\textsuperscript{119} By denying any “moral basis” for determining the character of prospective citizens, the two presidents promoted policies that would encourage immigrants to seek benefits from, or become dependent on, the administrative state.\textsuperscript{120} It became less important for immigrants to become naturalized citizens or to participate in the political process.\textsuperscript{121}

\textsuperscript{99} Id.

\textsuperscript{100} Id. at 137-38.

\textsuperscript{101} Id. at 139.

\textsuperscript{102} Id. at 139-40.

\textsuperscript{103} Id. at 141-44.

\textsuperscript{104} Id. at 87-124 (\textit{Progressivism, Immigration, and the Transformation of American Citizenship}).

\textsuperscript{105} Id. at 87.

\textsuperscript{106} Id. at 93.

\textsuperscript{107} Id. at 95 (quoting from \textit{Thomas G. West, Vindicating the Founders} 149 (1997)).

\textsuperscript{108} Id. at 96.

\textsuperscript{109} Id. at 98.

\textsuperscript{110} Id. at 99.

\textsuperscript{111} Id. at 100 (quoting \textit{Herbert Croly, The Promise of American Life} 418 (1911)).

\textsuperscript{112} Id. at 102 (quoting from \textit{John Dewey, The Ethics of Democracy} 6, 7, 13-14, 15 (1888)).

\textsuperscript{113} Id. at 106.

\textsuperscript{114} Id. at 109-10.

\textsuperscript{115} Id. at 113, 115-18.

\textsuperscript{116} Id. at 114 (quoting from \textit{Daniel J. Tichenor, Dividing Lines: The Politics of Immigration Control in America} 1-46-47 (2002)).

\textsuperscript{117} Id. at 120-22.

\textsuperscript{118} Id. at 120.

\textsuperscript{119} Id.

\textsuperscript{120} Id. at 120-21.

\textsuperscript{121} Id.
Marini concludes this chapter on a pessimistic note. In his opinion, the government's efforts to address the problems of citizenship and immigration have not lessened underlying conflicts about the protection of fundamental rights. The result has been a correlation of the rights of citizenship with the state or with "group identity." That places opponents of that result has been a correlation of the rights of citizenship with the state or with "group identity." That places opponents of that policy in jeopardy of being accused of racism. Accordingly, one legacy of the Progressive understanding of freedom is a modern state in which notions of race and class displace our original understanding of American citizenship that was based on an equality principle.

IV. REVERSING THE EXPANSION OF THE ADMINISTRATIVE STATE

Marini observes that in January 1981, President Ronald Reagan articulated a vigorous denial of the Progressive principle that a bureaucratic government could provide security consistent with the preservation of individual freedom. Reagan urged a course of action opposite to FDR's New Deal, questioning whether government by "an elite group" was superior to a tradition of self-governance and individual liberty. Marini credits Reagan with moving public sentiment against the excesses of big government and with reviving a public debate about the importance of limited government in maintaining a free society.

Reagan's concerns about the bureaucratic state and governance by elites foreshadowed—albeit in a different national and global context—the 2016 presidential campaign and the election of President Donald Trump. It is not surprising that Unmasking the Administrative State includes two essays that address our contemporary political situation. The 2016 election, Marini asserts, can be seen as a repudiation of the Progressive policies that have dominated both Democratic and Republican parties in domestic and foreign affairs since the end of the Cold War.

That repudiation, Marini contends, is also of the nation's governance by "professional elites" and a "policymaking establishment" based predominantly in Washington. The authority of intellectuals (liberal and conservative) previously had been unquestioned, and this was particularly pronounced in "official Washington," which had a critical stake in maintaining the status quo. But then-candidate Trump challenged even the intellectual authority of the leaders of organized conservatism. Trump also appealed to American citizens as citizens, not as members of discrete interest groups, and he also avoided making appeals to political leaders and political organizations. Marini attributes that latter strategy, in part, to Trump's recognition that political parties are weakly linked to the citizenry. In addition, Marini contends, the political parties themselves have too often agreed with the principles of the Progressives and their intellectual descendants. The notion of a common good had been eroded in favor of interest group and identity politics, but Trump rejected those categorizations. Trump's appeal, and electoral success, thus reflected the public's dissatisfaction with cultural transformations that occurred "almost completely outside the political process of mobilizing public opinion and political majorities."

Looking forward, Marini suggests some pathways by which the "centralization of politics, economics, administration, and public opinion" may be reversed or modified. First and foremost, the power of state and local governments must be restored. There also needs to be a revival of the "ground of politics" in the nation as a whole, an effort that could reinforce distinctions between "the social and the political," and the "public and the private," with a focus on reviving the institutions of civil society. Indispensable to this mission is a decentralization of authority from Washington, D.C.

Marini questions whether conservatism, properly understood, is simply an "antidote to liberalism." He goes on to present his own affirmative vision of conservatism. If conservatism means anything, Marini contends, it must "require a defense of the good as established by a tradition that has preserved the best of the past." That includes a defense of civil and religious liberty, founded in constitutional government. Marini urges conservatives to evaluate President Trump with reference to what he has accomplished since he has been in office.

---

122 Id. at 122-24.
123 Id.
124 Id. at 123.
125 Id.
126 Id. at 120.
127 Id. at 185.
128 Id. at 190.
129 Id. at 191.
130 Id. at 29-30.
131 Id. at 29-39 (written in July 2016), 273-86 (written before the 2018 midterm elections).
132 Id. at 274.
133 Id. at 34, 274.
as a whole," President Trump needs a "governing coalition" and Congress's cooperation.\footnote{149} This statement, written before the Democrats gained control of the House of Representatives in the 2018 midterm election, may be a statement of impossibility in the short term.

Marini, as a political scientist who has studied the rise of the administrative state over several decades, has succeeded in diagnosing fundamental problems of democracy and accountability in an agency-driven policymaking state. The central question that he poses is how to "restore the political rule of the people as a whole."\footnote{150} \textit{Unmasking the Administrative State} is an erudite, if sometimes esoteric, explanation of political philosophy. That is both a weakness and a strength. It is a strength because Marini offers a historical, systematic analysis of how our nation's constitutional thinking has evolved. But unfortunately, this book, in my judgment, does not provide a roadmap on how to restore our lost constitutionalism.\footnote{151}

This shortcoming may be the inevitable result of the problem of reconciling democratic processes, embodied in an elected Congress, with Congress's longstanding pattern of delegating its legislative powers to the presumed expertise of officials within cabinet agencies and so-called independent agencies.\footnote{152} Critics and reformers must continue to focus on tangible steps that can reconcile democratic governance with the need to address complex problems in our highly technological society.\footnote{153} Marini has identified one important change in direction—reallocating government power away from the central federal government to states and localities—which could be part of a practical solution to the centralizing of government functions.\footnote{154} It remains to be seen how other thinkers will supplement this solution with a more complete vision of renewed constitutionalism.

\begin{footnotes}

149 \textit{Id.} at 283-84.

150 \textit{Id.} at 284.

151 Unlike some other critics of the administrative state, Marini does not attempt to document specific failings or instances of overreach by agency officials. \textit{See Wallison} at 2-19.

152 \textit{See Gundy v. United States}, 139 S. Ct. 2116, 2129-30 (2019) (holding that a statute authorizing the Attorney General to specify the applicability of various statutory requirements to sex offenders convicted of offenses before the statute's enactment did not violate the non-delegation doctrine); \textit{see id.} at 2139-41 \textit{(Gorsuch, J., dissenting)} (criticizing Supreme Court jurisprudence that has upheld the constitutionality of statutory delegations of authority to agencies).

153 \textit{See The Federalist No. 37} (James Madison) ("Energy in government is essential to that security against external and internal danger and to that prompt and salutary execution of the laws which enter into the very definition of good government.").

154 \textit{See Jeffrey S. Sutton, §1 Imperfect Solutions, States and the Making of American Constitutional Law} 17 (2018) (explaining that state constitutional law "respects and honors" differences between and among States "by allowing interpretations of the fifty state constitutions to account for these differences in culture, geography, and history"); \textit{Id.} at 213-14 (emphasizing how our constitutional structure, including federalism, preserves liberty).
\end{footnotes}
An Interview with Professor John Pfaff of Fordham University School of Law

By Vikrant Reddy

In December, Congress passed and the president signed the First Step Act, the most important federal criminal justice reform legislation in a generation. The legislation made an immediate impact on the existing federal prison population by increasing the amount of good time credits that offenders can earn off their sentences (to a maximum of 54 days per year, up from 47), making retroactive the 2010 Fair Sentencing Act that reduced the disparities between federal crack and cocaine sentences, and even by including a provision to ban the shackling of pregnant women giving birth while incarcerated. The legislation is also expected to have long term effects through the allocation of $75 million over five years for expanded rehabilitative programming, a requirement that inmates be housed closer to their families to make visitation easier, and reducing the impact of federal sentencing enhancements such as the infamous § 851 (prior convictions) and § 924(c) (carrying a firearm).

Even after all of this, more legislation—a second step—seems inevitable. In April, President Trump announced that his administration was exploring ways to push even further on criminal justice reform. Moreover, new criminal justice legislation is one of the most frequently discussed topics among the candidates for the 2020 Democratic presidential nomination.

I sat down for a conversation about criminal justice reform with Professor John Pfaff of Fordham Law, the author of one of the most highly regarded criminal justice books of the decade: Locked In: The True Causes of Mass Incarceration—and How to Achieve Real Reform.

1. Your book, Locked In, is a critique of what you call the "Standard Story." What is the Standard Story?

   The “Standard Story” is a term I use to refer to three popular theories about the forces driving mass incarceration: increasingly longer sentences, the War on Drugs, and the power of private prisons. None of these theories is wrong—our sentences are long, the War on Drugs has sent a lot of people to prison, and private prisons generally are not advocates of reform—but each is far less important than most people think. The problem with the Standard Story is that focusing on these theories often blinds us to things that matter far more.

   Take long sentences. Compared to Europe, our sentences are long, and they did get slightly longer over the 1990s and 2000s. But the median time to release for someone sent to prison for a drug or property crime is just one year, and it’s barely four years for a crime of violence (with most very long terms being for homicide). Sentence length matters, but it turns out that
the number of admissions—and the role prosecutors play in setting those numbers—matters far more. Yet we still emphasize legislative reforms targeting sentence length while paying little (though increasing) attention to prosecutors.

Similarly, we focus a lot on drug crimes, even though only 15% of state prisoners are in for drugs, compared to 54% in for violence (and 90% of prisoners are in state prison, so they are far more significant than federal prisoners in our criminal justice system). Real reform requires us to change how we punish violence, particularly serious violence like homicide, but our emphasis on drug offenses has led the public to refuse to accept this. And while we talk a lot about private prisons, they hold only about 8% of all inmates, with 92% in public facilities. About two-thirds of the $50 billion we spend on corrections goes to wages and benefits for the public employees in those state-run prisons, yet by talking so much at the private prisons, we effectively give far more powerful organizations like correctional officer unions an undeserved pass.

2. Do current federal criminal justice reform proposals address your deeper critique of the criminal justice system, or do they misguided address the Standard Story? (Or neither?)

Things are a little different for the Feds. Due to some constitutional issues involving federalism, the federal system can only prosecute a limited number of cases—it’s really hard to face a federal arson charge unless you, say, burn down the White House. As a result, about 50% of all federal inmates are in for drugs (one of the few kinds of offenses the Feds can easily target), compared to 15% in the states. So the emphasis on drug sentencing in federal reform makes a lot more sense and is a lot more important.

Moreover, the power of prosecutors in the state system differs significantly from that in the federal system. Local prosecutors are far more autonomous, raising all sorts of complicated issues about how to regulate their decisions. In the state system, the prosecutors are almost always independently elected at the county level, while U.S. Attorneys are at-will employees of the Attorney General, who is the at-will employee of the President, creating a (possibly) more effective chain of command. That said, AGs have struggled to ensure consistency across the various U.S. Attorney offices, so independence and discretion still permeate the federal system.

Moreover, unlike the states, the Feds tend to have much longer sentences available, and are much more likely to impose those sentences on people convicted of drug offenses. So the Standard Story holds a bit more in the federal system.

Which actually points to one possible risk with federal sentencing reform. Even though the Feds hold about 10% of the nation's prisoners and vanishingly smaller shares of our jail detainees, parolees, and probationers, the federal system, and thus federal reform efforts, seem to receive about 95% of our media attention. This is one reason why the Standard Story has such traction: it fits the (distinctly idiosyncratic) Feds much more closely than the states. So while federal reform is worthwhile, the attention it receives runs the collateral risk of further convincing people that Standard Story-like reforms are appropriate for the states as well.

3. Many people think we could dramatically reduce incarceration and solve a host of other criminal justice problems simply by ending the War on Drugs—particularly because the trade in illegal drugs is intrinsically connected to violence. You think that’s too simplistic. Why?

To start, if we focus on just those who are in prison for drugs, it’s only 15% in the state systems. Were we to free every person in prison on a drug charge tomorrow, we’d still have 1.2 million people in prison and something close to, if not actually, the highest incarceration rate in the world.

Of course, that’s a narrow definition of the War on Drugs. Someone who commits murder in a drug deal gone bad is classified as serving time for “violence,” not “drugs,” even though the root cause of that murder was prohibition. So legalization would likely reduce some of those crimes, and thus some of those incarcerations. Sort of. Maybe. To some degree.

There are a few complications. To start, some, and perhaps many (though surely not all), of the homicides that arise from prohibition may occur in a world with legalization too, just for different reasons. Drug-related violence tends to occur in neighborhoods that are already under immense economic, and thus social and emotional, strain, which are all causes of violence; ending criminal prohibition may help, but if the underlying stress remains, so too will much of the violence. The nominal cause of the violence may shift (from a drug deal gone bad to some other dispute), but the levels may not change as much as many hope.

Moreover, legalization will lead to lower prices and less stigmatization, and thus more people using drugs (at least for some types drugs). For some, the cheaper drugs will mean they are more affordable, and will lead to less property crime to fund consumption. For others, the cheaper drugs will lead to increased use that further destabilizes the person’s life and may lead to more property crime. The effects are complicated and confusing and hard to fully predict in advance.

Will a shift to a legalized regime of some sort make things better? My guess, and I think that of most other people, is yes: that the net effect of legalization will be a net reduction in violent and property crime. But the dynamics will be complicated, and if we legalize without taking steps to address the public health issues of drug use and abuse as well as the other social pressures that lead to violence, the gains will likely be less than many people expect.

4. You are critical of the argument that private prisons drive increases in incarceration. Why? Incentives matter in political institutions. Isn’t it obvious that those who profit off incarceration would favor higher levels of incarceration? On a related note, didn’t an Obama Administration report conclude that private prisons are less well-run than public facilities?

To start, the data on the relative quality of public v. private prisons is quite thin, and people who study the issue rigorously note that what little data we have is wholly inconclusive. There are terrible private prisons and terrible public ones, and there are better private prisons and better public ones.

Here’s a confounding example. Until recently, the Florida Department of Corrections website proudly bragged that as a general matter its prisons did not have air conditioning—a feature that has led to inmate deaths from heat exposure in Florida and
elsewhere. But that same webpage grudgingly acknowledged that all private prisons in Florida were fully air conditioned. It’s complicated.

More important, only 8% of all prisoners are held in private prisons, most of those in just five states, and there’s no evidence that trends in those five states differ from those elsewhere.

You’re right that incentives matter, but (1) public prisons often profit just like private ones do, and (2) public prisons create “profits” in all sorts of other ways besides rewarding shareholders. To start, at least in large urban counties, the offices are staffed. And in most states, politicians with prisons in their districts gain because those prisoners count as living in that prison for legislative districting, even though they cannot vote (which, because prisoners are disproportionately people of color from urban places and prisons are mostly located in more rural areas, provides a clear electoral benefit to the Republican Party).

In the end, it seems likely that the availability of these sorts of public-sector benefits create incentives for public-sector resistance to reform that is far more significant and effective than private-sector resistance.

5. Your key argument has always been that the rise in incarceration has been caused by charging decisions made by prosecutors in recent decades. What kinds of reforms could address that problem?

We’ve started to see efforts, at least in urban counties, to elect reform-minded prosecutors, which I think is a good first step. But it is also important to note the limitations of elections-driven reforms. To start, at least in large urban counties, the offices are quite large, so reformist district attorneys may still struggle to get their message down to the rank-and-file prosecutors who make day-to-day decisions. Perhaps more concerning, it is easy now to sound like a reformer without having to necessarily be one: there’s a set of terms that are easily bandied about—“bail reform,” “low-level drug offenses”—but which are generic enough that they may not reflect much of a real desire for change, and that allow reform-sounding candidates to easily carve out exceptions that drown the rule once in office.

Given that it is increasingly easy to sound like a reformer without necessarily being one, I am glad to also see the rise of groups dedicated to making sure district attorneys’ actions match their words. In particular, we are starting to see more “court watching” groups. In New York City, for example, there’s a group called CourtwatchNYC, whose members sit in court daily all across the city publicly reporting on whether line prosecutors are upholding their bosses’ campaign promises. Twitter has also given public defenders a platform to report on whether actions track rhetoric as well.

I also favor the development of charging and plea bargain guidelines. This could curtail some of the abuses of discretion that prosecutors sometimes engage in, but more broadly it could address a problem with experience. We often call on newly-hired prosecutors fresh out of law school to make truly life-changing (for the defendant) decisions, like whether to set a reasonable bail. Guidelines could not only help minimize abuses, but they could simply provide prosecutors with a sense of what the optimal decision should be. So much of a prosecutor’s job is making complex risk assessments that law school did not train them to make; some sort of assistance is clearly necessary.

6. Didn’t higher levels of incarceration over the past twenty years correlate with a crime decline? This is the argument made by criminal justice reform skeptics like former Attorney General Sessions.

Sessions’ view is wrong, but it is important to understand precisely why. During the 1970s and 1980s, when crime was high and rising and prison populations were low and rising, prison growth probably did reduce crime; you can’t add 1.1 million people to prison and have no impact.

But that does not mean that prison was the right response to rising crime (although it may have been politically necessary, and it was certainly politically expedient). Prison worked, but other options—at the very least, increased policing—would have worked just as well at far lower social cost. If you get an infection in your finger and cut off your arm, you stop the spread of the infection. The amputation worked. But . . . maybe try an antibiotic first? Prison growth was a blunt tool with massive—and avoidable—collateral consequences.

And today? Sessions’ approach is even more invalid. It’s clear that in a time like now—with low crime and high levels of imprisonment—increased incarceration has close to no additional impact on crime. Evidence indicates that its deterrent effect is weak to negligible, and one impressive recent paper suggests that the more time someone spends in prison—the longer he is incapacitated and thus not committing crimes outside—the greater the risk of his reoffending upon release, to the point that the increased risk of reoffending almost wholly offsets the incapacitation effect.

7. You think prosecutors should be charging fewer people in ways that trigger prison sentences. You also argue, however, that most people in prison are there for violent acts. This seems like a contradiction. After all, isn’t it reasonable to expect that all people who have committed an act of violence serve at least some prison time?

Not necessarily. To start, it’s worth noting that lots of people who commit violent crimes don’t go to prison already: we arrest 500,000 people for serious violent crimes every year, but we admit fewer than 200,000 to prison annually for such offenses. My guess is that even more could be better served—solely from a public safety perspective—by options outside of prison.

Moreover, resources are limited, so a dollar spent on prisons is a dollar not spent on policing, or on non-police interventions like Care Violence, or on drug treatment. And as I just mentioned, prison is a relatively ineffective way to combat crime. So we could achieve the same reduction in violence at far less social cost by relying on non-prison approaches.

Moreover, any analysis of the gains from prison has to wrestle with its costs—not just the $50 billion we spend to run the prisons, but the surely far faster (but completely unmeasured)
social costs: the risk of physical and sexual assault in prison, the loss of income upon release, the financial and emotional and social costs to the inmate’s family and loved ones, the way in which prisons spread diseases and increase the risk of drug overdose deaths, and so on. Trying to get a rough handle on these costs is my next big project, and my guess is that the number is going to be staggering.

8. Do you believe, like many criminal justice reformers, that there is a “crisis” in indigent defense? If so, what policies would fix it?

The crisis is this: about 80% of all defendants facing prison or jail time count as indigent and qualify for a state-provided defense lawyer, and we grossly underfund their defense. At least as of 2006-2007—the last years for which the Bureau of Justice Statistics and the American Bar Association have comprehensive data—we spent about $6 billion on prosecution and about $4.5 billion on indigent defense. But that significantly understates the difference. Prosecutors’ offices have access to all sorts of free services that defense lawyers have to pay for: police departments to do the investigation and forensics, state labs to test DNA and alleged drugs. A study in North Carolina, for example, found that while the nominal budgets for public defense and prosecution were roughly identical, the prosecutors’ budgets were essentially triple the defenders’ once all the free services that prosecutors (alone) received were accounted for.

The easiest solution would be some sort of federal assistance for indigent defense. $4.5 billion is a rounding error in the federal government’s $3.5 trillion budget—federal grants could easily double or triple what we spend on defense for the poor. And this is an issue that is increasingly getting bipartisan support. The support from the left has always been there (at least nominally), but a recent article in the Marshall Project pointed to growing conservative support motivated by Second Amendment concerns: poor people are losing gun rights due to inadequate representation. More broadly, though, I would think that indigent defense should be a core conservative value: what is more conservative than taking steps to ensure that the weakest among us are protected from that state at its most powerful?

And if Congress won’t act, then perhaps the Supreme Court can. The thing about Gideon v. Wainwright—the case that requires states to provide the poor with lawyers—is that it is the classic example of a Supreme Court unfunded mandate: the Court tells states to provide lawyers to the poor, but says nothing about how. A more robust, substantive take on what counts as adequate defense—not at the individual case level, but at the more systemic level—could push states in better directions as well.

9. What is the moral hazard problem in our criminal justice system? How was that problem solved by California’s Realignment? On a related note, has Realignment caused crime to increase as some people worried?

The moral hazard problem that concerns me (and others) arises from the baffling and poorly-reasoned (if reasoned at all) way that we fracture political and financial responsibility for crime control across city, county, and state (and, to a much lesser degree, federal) jurisdictions. One such example involves the political and financial costs that prosecutors face. In almost all states, prosecutors are elected by county voters and mostly funded by county budgets. Jails—where we detain people pre-trial and for low-level misdemeanor offenses—are county-funded as well, as are many probation services. But prison, the punishment for serious felony offenses? That’s funded at the state level.

Look at what that does. If the prosecutor is more lenient and seeks a misdemeanor charge, his home county has to pay for the jail term. But if he’s harsher, not only does he look tougher on crime—a move that, more often than not, is politically safer—but it is fiscally cheaper for him, since a different government picks up the tab. Even if prosecutors aren’t cynically taking conscious advantage of “free punishment,” the moral hazard problem means they will ignore the costs of sending people to prison, which is the classic definition of an externality or moral hazard problem.

In response to a judicial finding that its prison system was unconstitutionally dangerous and overcrowded, California adopted a sprawling, complex reform law colloquially called “Realignment,” which has been quite successful in scaling back California’s reliance on prisons—about 45% of the national decline in prison populations since 2010 is just California’s decline. One major reason for this success is that Realignment insisted that county jails, not state prisons, had to house people convicted of a wide range of felonies, not just misdemeanors. In other words, it pushed the costs back onto the counties, and the counties balked. This is not the only reason for California’s decline—like I said, Realignment is complex—but it certainly appears to have mattered.

As for the rising crime some feared (and which some still tout), there is simply no evidence that it has actually happened. One study found a slight increase in auto theft, but a later study identified the same increase . . . and then found that it had subsequently disappeared.

The people pushing the “rising crime” story are those most politically opposed to reform: the police, sheriffs, and prosecutors. One of my biggest hopes about the current reform push is that we come to view these criminal justice actors as the political actors that they are—that we stop accepting their claims at face value and subject them to the same scrutiny and skepticism that we apply to those on the reform side.

10. Why does Chicago struggle with violent crime in a way that New York City does not? New York has done several things which reform opponents have worried will lead to higher crime—incarcerating fewer people, eliminating stop-and-frisk, etc.—but it continues to be one of the safest big cities in the country. Why can’t Chicago just be New York?

I think that’s a tough question for which no one yet has a good answer. I’d point out just three things, though. First, the focus on Chicago is peculiar. While its homicide rate is still higher than it was in the past, even at its recent peak its rate was lower than in other cities that receive no attention at all. And over the past two years, its rate has dropped significantly, yet it still remains our go-to example of “rising homicide.” Chicago is being singled out for reasons that are, sadly if predictably, unrelated to crime trends.

Second, it is worth noting that the year homicides spiked, about half the citywide increase in homicides occurred in just
five neighborhoods that were home to less than 10% of the city’s population. So the story in Chicago—and the story of crime in cities nationwide—is significantly local, even within the city, and thus defies any sort of easy “problem in Chicago” answer.

And third, while the causes of Chicago’s homicide spike, and broader gun violence spike, are complicated, one factor that certainly played some role was the state’s decision to defund Chicago’s Cure Violence program (called CeaseFire in Chicago), a street-level violence intervention program that relies on local community members to try to intervene to prevent retaliations after a shooting. Shootings and homicides in Chicago had been declining prior to the defunding and rose shortly after the state cut funds. More interestingly, one neighborhood’s program was spared the cuts, and that neighborhood did not see shootings or homicides rise, even as they rose in other areas whose programs had been cut.

It’s important not to oversell this story: the timing isn’t a perfect fit, and trends in complex social problems like homicide rarely if ever have a single neat explanation. But it seems likely that CeaseFire was playing a real role in Chicago, which suggests that we need not unleash on Chicago a heavy-handed police response.

Refreshing Candor, Useful Data, and a Dog’s Breakfast of Proposals: A Review of Locked In by John Pfaff

by Kent Scheidegger

John Pfaff gives us two books under one cover in Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform.1 In the first book, he tells us that nearly everything we have been told about so-called mass incarceration by his fellow “reform” advocates is false. His candor is a breath of fresh air. He convincingly makes the case with a mound of useful data.

The second book, in contrast, is thinly supported and heavily influenced by Pfaff’s predispositions. He tells us that high incarceration rates are caused primarily by overcharging prosecutors, though his data do not rule out alternative hypotheses. He claims that the election of tough prosecutors is caused by the “low-information, high salience electorate,” not by informed people who genuinely and justifiably disagree with him on priorities. The primary ingredients in his stew of solutions are tools to save the ignorant masses from themselves by making our society less democratic and our criminal justice decision-makers less responsible to the people. Other intriguing possibilities raised by his data go unexplored.

Pfaff does not define what he means by “reform,” but he appears to use that term for policies that have the single-minded purpose of reducing the number of people incarcerated. Obviously, that is not the sole or universally accepted meaning of the term in criminal justice. The Sentencing Reform Act of 19842 definitely did not have that purpose. In this review, I will put the word “reform” in quotation marks when used in Pfaff’s sense.

I. The (False) Standard Story

In the first half of the book, Pfaff describes what he calls the Standard Story and proceeds to demolish it. The Standard Story is the one that nearly all advocates of “reform” use to sell their proposals to legislatures, courts, and the public. The Standard Story’s central premise is that America’s high incarceration rate is caused by sentencing harmless, low-level, nonviolent offenders—especially those whose only crime was possession of illegal drugs—to long prison terms.

A. Causes of High Incarceration Rates

In a well-publicized book provocatively titled The New Jim Crow: Mass Incarceration in the Age of Colorblindness, Michelle Alexander claims that drug convictions account for the majority of the increase in prison population. President Obama picked up the theme in a speech to the NAACP, declaring that locking up nonviolent drug offenders is “the real reason our prison population is so high.”3

But it’s not true, as Pfaff demonstrates. Most of America’s prison population is in state prison, not federal. While the number of state prisoners convicted of drug offenses did indeed increase

---

1 Hereinafter “Pfaff.”
3 Pfaff at 21.
sharply from 1980 to 2010, the number is still not enough to be
the primary cause of the increase. Far more people are in prison
for violent crimes than are in for drug crimes. Of those who are
in prison on drug charges, only a small portion are genuinely
low-level offenders. In addition, people who have committed
both drug and violent crimes are often convicted only of the
drug crimes because those are easier to prove. The notion that
our prisons are full of low-level, nonviolent drug offenders is a
myth. Decriminalizing or even legalizing drugs would produce
only a modest drop in prison population.

So if we are imprisoning mainly violent offenders, and the
prison population is exploding, then one might think excessively
long sentences are the problem. Many people do. But this, too, is
wrong, Pfaff says. Despite some well-known anecdotal examples,
actual time in prison for the typical offender is much shorter
than most people think. For example, the median armed robber
convicted in 2010 was released in less than three years.4 Even
nominal life sentences do not generally result in offenders actually
spending their whole lives in prison.

The real reason the prison population has grown and
remained high despite falling crime rates, Pfaff contends, is that
a growing percentage of those arrested are being charged with
felonies rather than misdemeanors. Crime rates, clearance rates
(percentage of crimes solved), and arrest rates all fell between 1994
and 2008, but the number of felony cases filed rose substantially.
The chance of a filed case resulting in prison time has remained
stable.

The percentage of arrests resulting in felony charges is
therefore the only factor that increased in the system. Breaking
the numbers down by offense, it appears that this increase is
mostly for violent crimes. The prison admission/arrest ratio rose
substantially between 1991 and 2011 for murder/manslaughter,
robbery, and aggravated assault. It rose less for burglary, and not
at all for theft and drug possession. The ratio for drug trafficking
rose somewhat, but then declined slightly.5

Pfaff notes, correctly, that the data do not tell us the reason
for the increase in admissions per arrest for violent felonies.
“Maybe police are doing a better job investigating certain types of
crimes, maybe prosecutors are being more aggressive in charging
them, maybe judges are more willing (or compelled) to send
defendants convicted of them to prison.”6 Yet having set out three
plausible hypotheses, Pfaff spends most of the rest of book focused
on only one of the three: aggressive prosecutors. The police and
judges get only passing mention.

The fact that the admission/arrest ratio for murder and
manslaughter increased alongside the ratios for other violent
offenses suggests that the quality of arrests coming to the
prosecutors from the police may indeed be a major factor. These
numbers are primarily for murder and voluntary manslaughter;

 involuntary manslaughter cases constitute only a small fraction.7
It would be a rare case of intentional homicide where prison was
not the appropriate disposition. If only half of homicide arrests
in 1991 resulted in a prison admission, the most likely reason is
that a large portion of the cases had shaky evidence, resulting in
a dismissal or a generous plea bargain; more arrests today might
indicate that police are doing a better job of collecting evidence to
support their arrests. Yet Pfaff spends little time on this intriguing
hypothesis.

Pfaff spends half a chapter debunking the notion that
lobbying by private prison interests is the cause of high levels
of incarceration. He notes that the power of the commercial
corrections lobby “is overhyped at every turn.”8 There is no need
to go into that in any depth here. Private prisons are such a small
portion of the total—about 8%—that it would take anti-capitalist
tunnel vision to consider such an argument remotely plausible.

B. Costs and Benefits

The discussion of public finances is more salient. The period
of expanding prison population was a period of economic growth
and ballooning government budgets. While corrections spending
did rise in overall dollars, its rise as a percentage of state and local
budgets was modest during the rising crime years. With a lag of
a few years, the percentage fell after crime rates fell.

The notion that we are spending huge amounts on prison
and that this spending is crowding out other priorities such as
education, health care, and transportation is simply wrong. Pfaff
notes, “we don't really spend that much on corrections,” about
3% of total state budgets.9 This is heresy among many self-styled
reformers. It is also true.10

The main purpose of this line of argument is political. It
serves to win over fiscally conservative voters who would otherwise
be inclined to support tough sentencing, and it provides cover
for conservative politicians who might actually want to support
softer policies for other reasons.11 Politically, then, just as the
prosperity of the 1990s and most of the 2000s enabled growth of
prison populations, so the fiscal crisis of 2008 and the pinch on
government budgets provided a window for a facially plausible
argument that we could save money and relieve the pinch by
incarcerating fewer criminals. This argument has struck a chord
with many moderate and conservative voters.

Yet Pfaff worries that the window may be closing with
renewed prosperity. Although he doesn't say so, the window might
also close when the false claims of the Standard Story come to

4 Pfaff at 56 tbl. 2.1, with data from the Bureau of Justice Statistics, Data
Collection: National Corrections Reporting Program.
5 Pfaff at 74 tbl. 2.2.
6 Id. at 73.
light. Pfaff warns his fellow “reformers” that focusing on factors that are not really major contributors to the prison population risks postponing the “reforms” that actually would produce a major reduction in prison population (i.e., releasing more violent criminals) until after the window has closed.

The “reforms” undertaken so far—mostly focused on nonviolent criminals and largely excluding those convicted of violent crimes—have yielded far less fruit than is commonly believed, Pfaff contends. The much-heralded legislation in Mississippi, for example, merely brought that state’s incarceration rate down from a level far higher than the national average to a level substantially higher than the national average. After the initial drop, it leveled off. Pfaff says “short-run declines followed by stasis” is the typical result of reform legislation.

The national prison population fell four percent from 2010 to 2014, but two-thirds of that drop was in California alone due to developments there that Pfaff calls “highly idiosyncratic.” For the rest of the country, the drop has been modest. Moreover, the drop has not been solely due to “reforms,” perhaps not even primarily so. Pfaff provides a graph of annual changes in prison population growth, and it shows a point of inflection in the mid-1990s, shortly after the peak in the crime rate and long before any significant “reform” legislation. Prison population grows more slowly after that, with the rate of change finally going negative in 2010. The reductions since 2010, Pfaff says, are “not entirely attributable to reforms, but are partially the continuation of a preexisting trend tied to falling crime.”

The Standard Story has focused on nonviolent criminals, Pfaff notes, because “reform” advocates believe that is the limit of what is presently politically feasible. He acknowledges that is likely true, but he has a blind spot that lets him see only half of why it is true. He notes, correctly, that “reforms” cutting imprisonment for violent criminals would risk increasing violent crime, which would undercut political support for the whole effort. This is true, but not the whole truth. The other half is simple justice, a concept that is nearly absent from this book.

Imagine you are the victim of a brutal rape, and a year later you sit down at a restaurant and see the rapist at the next table, happily enjoying his meal and laughing with his friends. Even if we had a magic pill that 100% guaranteed he would never commit another sexual assault (we don’t), it would still be horribly wrong for him to get off lightly for such an evil act. As discussed earlier, Pfaff’s own numbers show that sentences for violent crimes are generally not greater than the perpetrators deserve. People value justice for its own sake, irrespective of the practical consequences of punishment. That is an independent reason for punishing serious crime, and it is an independent reason why any effort to reduce punishment for violent crimes faces stiff headwinds. Pfaff seems to be nearly oblivious to it. In a footnote near the end of the book, he finally acknowledges the issue, but he merely says it is “beyond the scope of this book.” That is a strange statement. The book is all about the punishment of crime, and the reasons why people vote for “tough on crime” prosecutors and legislators is a major theme. Yet one of the primary reasons they do so—because they think violent criminals deserve tough punishments—is “beyond the scope?”

On the purely utilitarian side, the costs and benefits of incarceration are complex, difficult to measure, and partly intangible. The most obvious benefit, of course, is reduced crime. Here, Pfaff is again more candid than many “reformers.” He acknowledges that statistical problems cause many studies to underestimate the utility of imprisonment in reducing crime. The stronger studies tend to show a greater effect. Steven Levitt’s estimate that high incarceration rates produced a quarter of the great crime drop of the 1990s is consistent with other work. That is a massive reduction in victimization and suffering. Pfaff asserts that the same decline could have been produced “by investing in other, less costly, less brute-force solutions.” Later in the chapter he suggests that hiring more police officers is one such solution.

Pfaff offers empirical support for the thesis that there are diminishing returns to incarceration. If that increased incarceration is caused by harsher sentencing of less serious offenders, it does make sense that there would be diminishing returns. It does follow, as he says, that when incarceration rates are high due to harsh sentencing, the rate can be brought down without a large increase in crime. Yet that is the “low-hanging fruit” that has already been picked, i.e., the “low-level offenders” who never did make up a large portion of the prison population.

Pfaff acknowledges that further cuts may reach a point where they cause increased crime, undercutting political support for further cuts. But because crime rates are affected by a great many factors and many of the other factors point to continued declines, Pfaff believes that we have a lot further to go before sentencing strictness drops below politically acceptable levels. However, California may have already reached that point. The 2020 ballot will test that possibility with an initiative to partially roll back some of the state’s recent “reform” measures.

As difficult as the crime-prison connection is to gauge, other elements of cost-benefit analysis related to possible changes in  

---

12 Id. at 110.
13 Id. at 111 fig. 4.1.
14 Id. at 112.
16 Pfaff at 287 n.7.
17 Several times over the years, Gallup has asked people on both sides of the death penalty debate open-ended questions about the reasons for their positions. Support for the death penalty may be considered the ultimate “tough on crime” stance. “Just deserts” types of reasons are regularly cited more often thanUtilitarian reasons. See Art Swift, Americans: “Eye for an Eye” Top Reason for Death Penalty, Gallup, Oct. 23, 2014, https://news.gallup.com/poll/178799/americans-eye-eye-top-reason-death-penalty.aspx.
18 Pfaff at 114.
19 Id.; see also Barry Latzer, The Rise and Fall of Violent Crime in America 232 (2016).
policy are even more nebulous, with little or no hard data. An absence of empirical evidence provides greater leeway to indulge one’s preconceived notions. Pfaff speculates about the benefit to 100,000 children who presently have one parent in prison instead of both parents at home. Wouldn’t releasing their parents be worth a five percent increase in aggravated assaults, and even some increase in the number of murders? Isn’t this a debate worth having?, he asks.

We can debate it, but I think the answer is quite obviously no. On the benefit side, Pfaff simply assumes that the incarcerated parents would remain in their children’s lives and that their influence would be positive. Yet he established earlier in the book that the notion of large numbers of low-level, nonviolent offenders is a myth, so we are talking mostly about parents who have committed serious crimes of violence. Pfaff offers no support at all for his assumption that these violent men (mostly) would be good fathers. I think of Huckleberry Finn. Huck suffered from the lack of a caring, nurturing, supportive parent, but he suffered a lot more when the drunken, abusive father he actually had showed up. For the dubious benefit of 100,000 criminal parents returning to homes where they may do more harm than good, Pfaff would accept about 40,000 people being attacked and severely injured, and some unstated number being killed. In my view, it is not a close question.

Pfaff notes as costs of incarceration the violence within prison, the loss of employment skills, and barriers to employment after release. These are inevitable costs to some extent, but not to the extent that they exist today. More effective prison discipline, better smuggling control, and greater employment of inmates (including unrestricted sale of prison-made goods in market segments with no substantial domestic production) are all matters that warrant action but will require overcoming political opposition. Employer reluctance to hire is more a consequence of the conviction than the sentence, but critical examination of government licensing requirements and encouragement of voluntary hiring of released prisoners in appropriate jobs are also worthwhile endeavors.

Pfaff ends the first half of the book with the conclusion that current “reform” policies are not reducing prison population as quickly as the “reformers” had hoped they would. He considers “faith in the Standard Story” to be one reason for this. Was it really faith, or was the Standard Story a deliberate fraud from the beginning? On the other side, he claims it is “fear”—not genuine, well-founded concern for the innocent victims of crime—that motivates opposition to softer sentencing practices. The second half of the book is devoted to ways to sharply reduce incarceration.

II. A (Dubious) New Narrative

The second half of the book is markedly more political and ideological and less data-driven than the first half. The overall thrust is that American criminal justice is suffering from too much democracy, and that the path to deeper “reform” is to take justice decisions away from the Great Unwashed and assign them to wise Philosopher Kings.

A. Supposedly Harsh Prosecutors

Pfaff argues that prosecutors have much greater power to determine a defendant’s sentence than is generally understood, and that their power is excessive. The argument is thinly supported. Pfaff points to mandatory minimums and uses the federal system as an example. While debunking the Standard Story, he correctly noted that most prisoners are in state, not federal, prisons, and that the federal system is not typical of the states. But in his indictment of prosecutors, he commits the very error he had critiqued. To support the proposition that in “many states” the prosecutor’s charge can restrict the judge to a narrow sentencing range, he gives us no hard data but only a general observation that “[s]everal states use sentencing guidelines . . . .” How many? How rigid are the guidelines? Do they allow departures? He doesn’t say.

Pfaff conceded that in his home state of New York the prosecutor’s charge does not set a sentencing floor. That is also generally true in California, where the judge has broad power to dismiss allegations that would otherwise set a minimum punishment, except in the few instances where that power has been revoked by statute for a particular charge. Even the much-criticized Three Strikes law permits judges to “strike strikes” to avoid the law’s mandatory sentence, including over the prosecutor’s objection. If prosecutors use their charging discretion to lock judges into unjustly harsh sentences, the objection, it should not be difficult to demonstrate that sentencing laws in states comprising the bulk of the population give them that power, but no such support is offered.

Direct evidence of prosecutorial harshness, Pfaff notes, is unavailable due to a lack of data focused specifically on charging decisions. He goes on to speculate on reasons, yet some of the reasons he cites are not harshness at all, as most people would understand the term. One possible reason for prosecutors seeking prison terms for a greater percentage of felons is that more defendants today have long criminal histories than in the past. The peak crime years of the 1980s and 1990s produced a lot of long records. “America is the land of the second chance,” President Bush famously proclaimed in his 2004 State of the Union Address. He did not say that America is the land of the seventh chance. To the extent that prosecutors are seeking more severe punishments

23  Pfaff at 132.
24  Id. at 131.
25  Id.
26  CAL. PENAL CODE § 1385.
28  Pfaff at 134.
for habitual criminals than they do for first-time offenders, they are not acting harshly. They are doing their jobs exactly right. Yet Pfaff calls it “harsher.”

Another possibility, noted in the first part of the book, is that improved policing may be creating stronger cases against felons. Perhaps criminals were previously getting off with unjustly lenient sentences or walking altogether due to inadequate evidence, and with better police work and technology such as DNA and widespread video cameras they are now getting the punishment they deserve. If so, and if we care about justice, the change should be applauded, not lamented. That second “if,” however, seems to be one of Pfaff’s ideological blind spots. If he does care about sentences being just punishment for the crimes, he doesn’t show it in this book.

B. Too Much Democracy

In most states, prosecutors are elected locally, either by county or, in rural areas, by judicial districts with more than one county. Pfaff says the record on election of prosecutors is “bleak,” because nearly all are reelected and most are unopposed.

It is true that voters often know little about their local prosecutor, but that is not necessarily a problem. Absence of local news about the prosecutor’s office, except in cases of exceptional crimes, indicates a lack of controversy. Lack of controversy, in turn, indicates that the district attorney is conducting the office in a way that people would approve of. There is no shortage of lawyers, and even some judges, who would like to take that office for themselves if they thought they had a shot. The fact that incumbents are rarely challenged indicates that potential challengers do not think voters would find the incumbent’s policies objectionable if they were brought to their attention in a campaign. The fact that most incumbents are reelected when they are challenged tends to confirm that judgment.

What is bleak about that? Elected officials who pursue policies consistent with the people’s view of justice is exactly how democracy is supposed to work.

In the federal system, the people elect only the President, and all other executive policy-making officials, including U.S. Attorneys, are appointed by and can be fired by him. Is this unitary executive of the federal government better than the distributed executive of separately elected officials, including local prosecutors, that we have in most states? Pfaff thinks so, even though federal prosecutors have been widely criticized by others for being overly aggressive. The people of the states, as they have written and rewritten constitutions in the last 240 years, have chosen not to follow the federal model. They have instead made more officials directly answerable to the people. Should we cast aspersions because nearly all are reelected and most are unopposed.

Pfaff laments the fact that locally elected prosecutors are not constrained by the cost of imprisonment when they seek prison sentences because prison costs are paid at the state level. He considers this to be a “moral hazard.” Most people, I suspect, would consider determining punishments according to justice and not dollars to be a valuable feature of the system, not a problem to be fixed. Do we actually want a world where a prosecutor tells the victim, “I am only going to ask for two years in prison for the man who raped you because we can’t afford any more”? Do we want a person who was caught burglarizing a home and found in possession of stolen property from fifty unsolved home burglaries to get probation because the jurisdiction wants to cut prison costs? How many more people’s homes would be invaded as a result?

Pfaff notes with approval California’s “realignment” experiment, which places lower-tier felons in county jail rather than state prison. However, this change was not aimed at changing prosecutors’ charging decisions. It has produced a one-time drop in overall incarceration, to be sure, but Pfaff cites no evidence that changes in charging are the cause. He does cite a study by a left-leaning think tank claiming that realignment produced only a small increase in property crime “compared to what it otherwise would have been.” However, the study says, “we find robust evidence that realignment is related to increased property crime”; considering the source, that is quite an admission. Overall, California’s property crime rate ran about 9% below the national average in the years leading up to realignment and has run about 5% above the national average since, a 14% jump in relative rates. California is hardly a model for the rest of the country to emulate.

The root problem, in Pfaff’s view, is voters he regards as ignorant and inconvenient. He identifies several “defects” in the politics of crime. First, voters are more likely to punish a prosecutor or a judge at the polls for a “false negative”—a failure to sufficiently punish a person who should have been confined and subsequently commits a major crime—than for a “false positive”—an excessive sentence of a person who would not have committed a new crime even if he had not been incarcerated.

Pfaff supports his argument about the “false positive defect” with an atrocious misuse of the old maxim that it is better for ten guilty men to go free than for one innocent one to be convicted. He blithely assumes that the calibration of punishment for a person who is truly guilty belongs in the same tier of justice as

---

30 Pfaff at 136.
31 Id. at 138-139; see supra text accompanying note 6.
32 Pfaff at 141-142.
33 Id. at 144.
35 Id. at 142.
36 Id. at 150-151.
37 Id. at 152.
38 See id. The only citation he gives is to a follow-up study, but the initial study is evidently Lofstrom & Raphael, Public Safety Realignment and Crime Rates in California (PPIC 2013), https://www.ppic.org/publication/public-safety-realignment-and-crime-rates-in-california/
40 Pfaff at 167; see 4 William Blackstone, Commentaries 352 (1st ed. 1769).
the determination of whether he is guilty or innocent. While one might conceive of hypotheticals where overpunishment amounts to an injustice of the same magnitude as sending an innocent man to prison, there are no such examples in American law today.41

A more realistic example illustrates that the 10-to-1 maxim may work the other way around when it comes to sentencing. Suppose ten people are convicted of rape, and the question is whether to sentence them for three or six years.42 Let us further assume—generously—that nine of the ten would be law-abiding in years four to six, but one would commit another rape. Is it better that ten rapists serve an extra three years in prison for evil acts they chose to commit than that one innocent person be raped? It is in my book.

Incarcerating a convict for a period in which he would not have committed a crime is not a false positive because incapacitation is not the sole reason for imprisonment. Justice is a sufficient reason by itself. In my example, we should not wring our hands for the rapists who do six years because that is not greater than their just deserts for their crime, whether there is any utilitarian benefit to the additional time or not. The sentences Pfaff regards as false positives may well be sentences that the typical voter regards as just. Perhaps district attorneys who seek sentences for burglars that result in their serving a year and two months, on average,43 usually have no opponent, not because the people are ignorant of false positives, but because anyone who ran on a platform that such sentences are unjustly long would be greeted with derision and lose in a landslide in most jurisdictions.

The complement to the false positive problem is the idea that voters pay excessive attention to rare but extreme cases. The problem here, according to Pfaff, is “low-information, high-salience” voters who only know about the headline-making cases, not how the system normally works. Pfaff’s example of this phenomenon is revealing: the Willie Horton case. Pfaff recites the version of the Willie Horton fable that casts former governor and presidential candidate Michael Dukakis as the victim of a misleading smear. In this telling, Dukakis had simply inherited a prison furlough program from his predecessor. It was largely successful, but one isolated failure caused a public relations disaster. Horton absconded from the program and committed a brutal home invasion, aggravated assault, and repeated rapes. The Bush campaign, Pfaff says, “released a powerful attack ad with racist overtones . . . using the Horton case to argue that Dukakis was soft on crime.”44 Conspicuously absent in this telling is any mention of what crime Horton was in for when he was furloughed.

Horton was in for murder and ineligible for parole; a rational furlough program would not have included people with such sentences.45 True, Dukakis had inherited the program from his predecessor, but it was a court decision construing the statute that made murderers eligible, and Dukakis vetoed a bill to tighten up the system. He continued to resist even after Horton’s crime. Far from “racist overtones,” the Bush campaign’s ad refrained from using Horton’s picture or mentioning his race.46 The ad with the picture was produced by independent operatives, and the contemporary criticism of Mr. Bush was merely that it took his campaign too long to express disapproval.47

In short, the Bush campaign’s attack on Dukakis’s furlough program was a valid campaign challenge to an astonishingly ill-considered policy, one which had tragic consequences for innocent people. The fact that Pfaff uses the stock story without fact-checking is typical of the mindset that pervades this book. All who support “reform” get a pass, even when spreading or accepting disinformation, and all who oppose it get disparaged on thin or false evidence. Pfaff uses the Horton ad myth to brand accepting disinformation, and all who oppose it as “low-information, high-salience” (LIHS) voters. What about the voters who agree with Pfaff on “reform” because they bought the Standard Story after hearing about rare, atypical long sentences for relatively minor crimes? The LIHS label would apply just as much to them, but they do not receive it in this book.

Public reaction to atypical crimes (and sentences) is a problem to be sure, but Pfaff fails to make the case that this is a major reason why supposedly harsh prosecutors are so routinely reelected. The alternative hypothesis—that people are generally satisfied with their elected prosecutors’ overall charging practices—is not explored.

Comparing the United States and Europe, Pfaff notes that politicians generally are more sensitive to voters’ wishes in this country because our government is much more decentralized. Indeed, our federal system, separately elected legislative and executive branches, and variety of independently elected officers in state and local government produce a degree of accountability that is unique in the world. Pfaff recounts this aspect of American exceptionalism as if it were a bad thing. It produces a result he disagrees with, to be sure. Yet this extensive separation of powers was deliberately designed to preserve freedom from the dangers of excessively concentrated power. We should be very careful about tinkering with it just to change a result on one policy question.

C. The Third Rail: Releasing Violent Criminals

With the foundation built to this point, Pfaff advances his main theme. “If we are serious about wanting to scale back
incarceration, we need to start cutting back on locking up people for violent crimes.”

The third rail of an electric train system is the one that carries the electric power. It delivers high currents at high voltage, and therefore it would be instant death to touch it, at least if one does so while connected to ground. It has become a widespread metaphor for a political issue so dangerous to one’s reelection chances that no politician wants to touch it.49 In an attempt to lower the voltage, Pfaff argues that we do not need to incarcerate violent criminals to the extent we do for the purposes of incapacitation or deterrence; he simply refuses to discuss justice.50 Indeed, he goes so far as to assert that less punishment of violent crimes will actually make us safer.51 Unlike his thorough, data-grounded demolition of the Standard Story, however, the evidence cited here is off-target.

Pfaff notes that most offenders follow a life course in which they are more likely to commit crimes while relatively young and less likely to do so in their older years. He acknowledges that some people do not follow this course and do not desist.52 Therefore, he claims, “long sentences frequently over-incapacitate. We don’t need to lock up most violent twenty-year-olds for thirty years to keep ourselves safe, since most of them would naturally desist from offending much sooner than that.”53 Yet Pfaff’s own demolition of the Standard Story shows that this argument is off-target. We don’t need to lock up most twenty-something violent offenders for that long, and indeed we are not doing so at present. Just four pages earlier, Pfaff noted that violent offenders admitted in 2003 served an average of only 3.2 years.54 Pfaff points to data on people released after California revised its Three Strikes law for the proposition that this group committed fewer crimes after release “because they are simply older.”55 Indeed, the one treatment that we know works to decrease violent propensities is aging.56 So while we may not need to lock up a violent 24-year-old until he is 54, locking him up until he is 34 may indeed be quite effective in preventing a number of crimes of violence and sparing multiple innocent people from victimization.

Pfaff’s extended attack on long sentences is very odd in light of the fact that he so thoroughly documented in Chapter 2 that long sentences are not the cause of mass incarceration. Although very long terms make headlines, there are too few of them to be a major factor in the incarceration rate.57 A term of years in single digits is not an unjustly long one for a violent felony, and Pfaff fails to make the case that such terms are ineffective as incapacitation.

Pfaff also attacks the deterrence rationale for long sentences,58 and this attack is similar in that Pfaff again misses his own target—the rate of prison admissions—and instead hits the one he has shown is not important in the large scheme of things: very long sentences. He begins by noting that certainty of punishment is more important for deterrence than severity of punishment. That is true, but no one, to my knowledge, is arguing to the contrary, or has made such an argument for a very long time.59 The question is not whether severity has a greater effect than certainty, but simply whether severity has an effect that makes a significant contribution that is worth its cost. “Empirically,” Pfaff notes correctly, “it is hard to separate out the deterrent impact of longer sentences from their incapacitative effect . . . .”60 Yet he curiously fails to mention a well-known study by a well-known economist whom he cites elsewhere in the book, Steven Levitt.61 Levitt and Daniel Kessler used the changes in crime in the wake of the enactment of sentence enhancements in California’s 1982 Proposition 8 to make the separation between deterrent and incapacitative effects. Immediately after the law’s enactment, potential criminals would be facing increased time but would not have actually served any of the additional time. Kessler and Levitt found that there was indeed a significant drop in crime that cannot be attributed to incapacitation.62

This is a critically important point of the argument: whether reducing sentences for violent crime would cause increased victimization of innocent people through loss of deterrence. Yet Pfaff cites only a single article by Daniel Nagin for the proposition that there is “little or no evidence that long sentences have any real deterrent effect.”63 What Nagin actually says is that “there is little evidence that increases in the length of already long prison sentences yield general deterrent effects that are sufficiently large to justify their social and economic costs.”64 If we assume for the moment that Nagin is correct, this supports only the proposition

48 Pfaff at 185-186.
50 Pfaff at 190.
51 Id.
52 Pfaff at 191.
53 Id. at 192 (emphasis added).
54 Id. at 188.
55 Id. at 193.
56 That is to say that statistically the rates of violent offending tend to decline substantially with age. See Pfaff at 191. There are, of course, exceptions. The infamous Lawrence Singleton—who raped a teenage girl, cut off her arms with a hatchet, left her for dead, and was released after a shockingly short eight years—later murdered a woman at the age of 69. See Singleton v. State, 783 So. 2d 970, 973 (Fla. 2001).
57 See Pfaff at 52.
58 Id. at 193.
59 Pfaff cites Gary Becker’s pioneering work from a half a century ago for a claimed mathematical equivalence of severity and certainty. Id. at 193 & n.21. He does not offer any evidence at all that this old proposition is necessary for deterrence theory or that anyone in recent years has asserted its truth. In every field, the pioneers were wrong about some things. We can see farther because we stand on the shoulders of giants. See JOHN BARTLETT, BARTLETT’S FAMILIAR QUOTATIONS 313 (15th ed. 1980) (quoting Sir Isaac Newton).
60 Pfaff at 193.
61 See Pfaff at 114, citing Levitt on another point.
62 Daniel Kessler & Steven Levitt, Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation, 17 J. LAW & ECON. 343 (1999).
63 Pfaff at 193.
that laws like Three Strikes, which impose life sentences on felons who would otherwise receive, say, ten years, do not have enough deterrent effect to justify the cost of imprisonment. But Three Strikes was always about incapacitation, not deterrence, so this proves nothing at issue. It does not negate Kessler and Levitt's result that an additional five-year term imposed on a repeat robber or burglar who would otherwise be getting several years does have a deterrent effect.

Earlier in the book, Pfaff maintained that the primary engine of prison growth was prosecutors seeking prison time for more of those who committed violent felonies, but that those convicted did not serve excessively long terms. When he turns to his critique of criminal sentencing, though, he tries and fails to refute the proposition that long sentences have a deterrent effect and save innocent people from victimization.

When we examine this chapter critically, what we see is that releasing violent felons is a political third rail for good reason. Some excesses, such as the original version of California's Three Strikes, can be pruned back, but Pfaff fails to demonstrate that overall the status quo of prison sentences for violent crimes is not needed for incapacitation and deterrence, as well as just plain justice.

D. A Dog's Breakfast of Proposals

After following this long road of statistics and studies, where do we go next? The end of the book is an untidy and unsatisfying scattering of proposals, a dog's breakfast.65

1. Guidelines

To deal with the perceived problem of too much democracy, Pfaff advocates adoption of sentencing guidelines.66 Continuing with the third rail metaphor, one might say that sentencing commissions are offered as an insulated glove. Politicians can touch the third rail through the insulation of a law that creates a sentencing commission rather than directly reducing sentences, and the unelected commission does the unpopular work of releasing large numbers of violent felons.

However, whether a guidelines system increases or reduces sentences depends on how the system is designed, how the commission operates, and who is on it. Pfaff makes clear that whether a guidelines system is good or bad depends, in his view, entirely on whether it achieves the result he considers desirable: reduced sentences.67 "It comes down to design," he says, but he does not tell us anything about the design other than that the sentences should be shorter.68

2. Increased Parole for Violent Felons

Increased use of parole is also a possibility, but Pfaff notes that nearly all proposals to date have been limited to prisoners designated nonviolent, meaning those whose current offense of commitment is for a crime labeled "nonviolent."69 To expand parole to more violent felons while insulating legislators and parole officials from the third rail of a Horton-type incident, Pfaff advocates the use of "quantitative risk-assessment tools"—parole by computer algorithm.70 He deems the victimization of additional innocent people to be an acceptable cost of paroling more violent criminals, but he wants to insulate the people who make that value judgment from the political consequences.

Recent enactments suggest that most politicians do not believe that a program that forthrightly embraces greater parole opportunities for violent felons is politically viable. California's Proposition 57 was a proposal so soft on crime that Governor Jerry Brown did not think he could get it through even that state's very soft-on-crime legislature, so he resorted to an initiative, which passed in 2016. He marketed it to the people as limited to inmates "with nonviolent convictions" and with a promise that it "[k]eeps the most dangerous offenders locked up.”71 That marketing was dishonest in two ways. First, the state's statutory definition of "violent felony" was written for a limited purpose and leaves out a great many crimes that most people would consider violent, including rape of an unconscious person, assault with a deadly weapon, and drive-by shooting.72 Second, inmates with a present conviction for a nonviolent felony but a slew of violent priors are eligible. The fact that savvy politicians intend on softening criminal punishment considered the deception necessary indicates that a proposal expressly including violent felons would be unlikely to pass.

A further step in the direction Pfaff proposed, enacted after publication of his book, is the federal First Step Act, which provides a system of credits for early release from federal prison. That act has a list of 52 exclusions from eligibility for its credits program.73 It was also marketed as excluding "violent and high-risk criminals" from the time credits program,74 though the list, long as it is, is not comprehensive. The act also requires for eligibility a classification by the Bureau of Prisons as "low-risk," so its actual application to felons convicted of violent crimes remains to be seen.

Neither of these programs fully embraces Pfaff's proposal of expanded parole for violent felons, and both were marketed as rejecting it. Even with the cloaking device of parole-by-algorithm, it seems unlikely that his proposal will be able to get much


66 Pfaff at 196-198.

67 Id. at 198.

68 Id.
on recidivism rates. It is probably for this reason that the First no evidence that any of the programs had a significant impact studies that could be considered acceptable, there was virtually loss. I’m not holding my breath for either of these. That means the benefits of alternative programs really would be determined by strictly valid methods. It would also mean that getting even limited “reforms” enacted. Changing the conversation in the way he suggests might result in the “reform” movement grinding to a halt.

4. Requiring Cost-Benefit Analysis

Pfaff acknowledges that the non-incarceration alternative punishments he favors require greater initial cost outlay than imprisonment, so to make this go down easier he proposes use of broader cost-benefit analysis. One problem with that proposal is that it requires quantifying the benefit of the supposedly more effective alternatives to imprisonment. It is doubtful that many “reformers” actually want rigorous quantitative assessment of their alternatives. The last thing they want is a repeat of the debacle of the 1970s. At that time, a study of studies found that many studies of the effectiveness of rehabilitation programs lacked the basic elements of methodological validity. Evaluating the studies that could be considered acceptable, there was virtually no evidence that any of the programs had a significant impact on recidivism rates. It is probably for this reason that the First Step Act defines “evidence-based recidivism reduction program” without any requirement of methodological rigor whatever, throwing open the door to “junk science.”

I would be delighted to see cost-benefit analysis done right. That means the benefits of alternative programs really would be determined by strictly valid methods. It would also mean that the costs of crime are fully accounted for, not just direct, tangible losses. I’m not holding my breath for either of these.

5. Regulating the Prosecutor

The biggest set of proposals has to do with prosecutors because of Pfaff’s dubious assertion that prosecutors are the “main engines driving mass incarceration.” From this premise, he considers regulating them to be a top priority. Pfaff’s first proposal is to beef up funding for public defenders so that they can “regulate” prosecutors by doing a better job of opposing the prosecution. This is one proposal that most prosecutors, along with nearly everyone in the criminal justice system, would agree with. Prosecutors uniformly tell me that they prefer that the defendant be well represented. The opposition to better funding for public defenders comes from those seeking the same government dollars, so perhaps “reformers” and advocates of more effective law enforcement could make common cause in seeking a larger slice of the budget pie for criminal justice generally. The notion that better indigent defense would actually have a significant impact on imprisonment rates, though, is not supported by any real data. Most cases end in plea bargains. Would more money for defendants mean bargains with shorter sentences on average? Maybe a little, but the idea that it would make a substantial difference, and that the difference would be a good thing, requires firmer support than it receives here.

Pfaff also wants grant programs to “encourage prosecutors to focus on more serious offenses.” Does he think that they are not focusing on serious offenses now and that money is the reason? He offers only one example of one odd grant program from the previous century. Most prosecutors are motivated by justice, and justice cries most loudly in the most serious cases. Pfaff argues that higher murder clearance rates in richer than in poorer portions of Los Angeles County result from prosecutors focusing more on solving crimes where rich people are victimized, but he offers nothing but rank speculation to support this argument. The most likely reason is that witnesses are much more afraid to come forward in gritty gangland than in leafy suburbs, and they are right to be afraid. Pfaff has made no case at all that prosecutors’ focus is not presently where it should be, so a proposal to change it with grant money falls flat.

Pfaff proposes gathering more data on how prosecutors exercise their discretion, but he offers few specifics and acknowledges the difficulty and danger of quantifying such a

75 Pfaff at 199-200.
76 Id. at 205; see, e.g., supra note 69, Argument in Favor of Proposition 57 (“focuses resources on keeping dangerous criminals behind bars”).
78 Id. at 48-49.
79 18 U.S.C. § 3635(3). This was done over my vigorous but ineffective protest. See Kent Scheidegger, Faux Pas Act Up for Senate Vote, CRIME & CONSEQUENCES BLOG (Dec. 11, 2018), http://www.crimeandconsequences.com/crimblog/2018/12/faux-pas-act-up-for-senate-vot.html. I have been calling for better research into what works since my 2008 remarks to the U.S. Sentencing Commission Symposium on Alternatives to Incarceration, but there is curiously little interest.
80 Pfaff at 206.
81 Id. at 207.
82 Pfaff’s idea of federal funding for local indigent defense, id., however, is a non-starter with federalists. His suggestion that the defense be funded at 80% of the prosecution is both unrealistic and unnecessary. The prosecution carries the massive burden of proof beyond a reasonable doubt, while the defense need only raise a reasonable doubt.
84 Id. at 208.
85 That is why many prosecutors continue to seek death sentences for the worst murderers, despite the massive costs created by the Supreme Court’s chaotic case law on the subject.
86 Pfaff at 209.
complex exercise of discretion.\textsuperscript{87} If people are evaluated on the basis of what is easily quantifiable rather than what is really important, their efforts can be skewed in undesirable ways. Just as the No Child Left Behind Act resulted in “teaching to the test,” so data-driven evaluation of prosecutors could subvert justice in pursuit of numbers. The devil is in the details, and we do not have enough detail to evaluate this proposal.

The next proposal is guidelines for prosecutors. To the extent that Pfaff proposes that elected local district attorneys have their discretion controlled by the state attorney general,\textsuperscript{88} he is proposing doing away with the localization of authority and accountability that the people of all but a few states have found to be needed. Local control affords each person a greater voice in decisions that affect the quality of life in their community. It also means greater freedom for an individual to move away from a jurisdiction that makes bad choices, since it is easier to move to another city than it is to move to another state. These are not small matters to be given up lightly, and a much stronger case would need to be made before taking such drastic action.

To the extent that Pfaff proposes that a district attorney’s office create internal guidelines so that junior prosecutors are not “winging it” with major decisions about people’s lives,\textsuperscript{89} he is right. That work is already underway. It began, in many offices, with a formalized process for deciding whether to seek the death penalty in cases where it is legally available.\textsuperscript{90} A less formal, but still not completely untethered, process is appropriate for less serious cases. Small counties with only a handful of prosecutors might not develop their own guidelines, but they can use those from larger offices as a starting point.

Pfaff proposes tackling the issue of locally elected prosecutors making decisions which impose costs on the state. As discussed earlier,\textsuperscript{91} he calls this a “moral hazard,”\textsuperscript{92} but the fact that prosecutors do not consider costs when deciding how to charge is a feature, not a bug, of the current system. Justice should not be decided on a price tag. The notion that this results in funding prisons instead of other measures, such as more police, is based on an unrealistically compartmentalized notion of public finances. If more funding for police is needed—and I agree that it is—there are other places to get the money, as discussed further in the next part.

After proposing to make prosecution discretion less local with statewide guidelines, Pfaff flips and proposes to make it more local by electing prosecutors from smaller districts. Specifically, he thinks that suburban voters in counties containing both high-crime inner cities and lower-crime suburbs have too much influence, and that they disproportionately favor tougher crime policies. He offers no evidence for this hypothesis. He notes that there are suburb-free prosecutorial jurisdictions in Baltimore, St. Louis, and New York City, but he provides no evidence that these cities have less of the problem he perceives.\textsuperscript{93}

There is some danger in going too small with prosecutor offices. Particularly in places with extensive organized crime, the possibility of intimidation or bribery looms larger. Baltimore is the city that resembles Pfaff’s proposal most closely, and to say that it is not a model of excellent prosecutorial practice would be an understatement. Again, a much stronger case would need to be made before this proposal could be seriously considered.

As his final proposal for prosecutor regulation, Pfaff calls for activists to mobilize to vote out tough prosecutors. He cites the example of Kim Foxx ousting Anita Alvarez in Cook County (Chicago and vicinity) in 2015.\textsuperscript{94} However, his discussion of that election has a glaring omission. After putting arguably excessive emphasis on the role of money in his criticisms of prosecutors, Pfaff fails to mention the massive infusion of campaign cash by George Soros toward the election of softer prosecutors, including Foxx.\textsuperscript{95} It is a billionaire swapping relatively low-budget races to give one candidate an overwhelming funding advantage, much more than mobilized activists, that has elected these “reform” prosecutors. Notably, Soros hit a brick wall in California in 2018, after this book was published.\textsuperscript{96} One of his 2016 winners has already thrown in the towel on reelection,\textsuperscript{97} and Foxx is on shaky ground.\textsuperscript{98} This tide may be starting to turn the other direction.

6. Disenfranchising the People

If convincing voters to elect “reform” prosecutors is unfeasible, another alternative for the infinitely wiser people is to just cut them out of the decision process. Pfaff proposes that we appoint both judges and prosecutors instead of electing them, as is done in the federal system.\textsuperscript{99} The people still elect the appointing official at some point, of course, but as long as crime rates do not return to 1980s levels, elections will still be decided largely on other issues. Pfaff cites the recall of Santa Clara judge Aaron Persky for an appallingly lenient sentence for a rapist as

\textsuperscript{87} Id. at 210.
\textsuperscript{88} Id. at 211.
\textsuperscript{89} See id. at 212-213.
\textsuperscript{90} See e.g., Gregory Totten, The Solemnity of the District Attorney’s Decision to Seek Death, IACJ Journal, 45 (Summer 2008).
\textsuperscript{91} See supra note 35 and accompanying text.
\textsuperscript{92} Pfaff at 213.
\textsuperscript{93} Id. at 215-216.
\textsuperscript{94} Id. at 216.
\textsuperscript{96} See Dan Walters, Billionaire Soros was a big loser in California vote, CALmatters (June 27, 2018), https://calmatters.org/articles/commentary/george-soros-california-2018-primary/. The three Soros-backed challengers all lost, and the sole Soros-backed incumbent barely won in a type of race where incumbents normally win handily.
\textsuperscript{97} Monivette Cordeiro & Jeff Weiner, Aramis Ayala won’t seek re-election as Orange-Osceola state attorney; Belvin Perry may enter race, Orlando Sentinel (May 28, 2019, 5:40 PM), https://www.orlandosentinel.com/news/breaking-news/os-ne-aramis-ayala-no-re-election-run-orange-osceola-state-attorney-20190528-z65rv7r84jd7wo6jpf06juu-story.html.
\textsuperscript{99} See Pfaff at 217.
an example of why elections for judges are a bad idea.\textsuperscript{100} In so
doing, he commits a fallacy that he denounces elsewhere in the
book: citing an example that is newsworthy because it is so rare,
but treating it as an instance of a widespread problem. California
judges are almost never recalled, and they are not commonly
opposed in their regular re-election bids. He notes that while there
is some movement to eliminate elections for judges, there is almost
none to eliminate elections for state and local prosecutors.\textsuperscript{101}

If the people of almost all of our states have settled on the
same system of locally electing prosecutors, there must be a good
reason for it. Yet if Pfaff understands the reasons that counsel
against his proposal, he gives us no sign of it. Decentralization and
direct accountability to the people are important elements of our
system of government. Centralization of power gives rise to a host
of evils. One would need a powerful case to overturn the collective
wisdom of our nearly unanimous states, and mere disagreement
with charging policy does not come close to that mark.

As for sentencing laws, Pfaff fortunately does not go so
far as to suggest appointment of legislators. Instead, he suggests
insulating the legislature from the political consequences of
subordinating public safety to other goals by outsourcing the
hard decisions on sentencing to a sentencing commission. Pfaff is
commendably candid in admitting his anti-democratic motivation
for supporting commissions. He quite forthrightly (one might
even say blatantly) defines the “success” and “effectiveness” of
sentencing commissions solely in terms of whether they reduce
prison populations, not whether they make the sentencing systems
of their jurisdictions fairer, with sentences proportioned to actual
culpability.

This is not to say that sentencing commissions are
necessarily a bad idea. The problem with making criminal law by
legislation is not excessive democracy, or even lack of expertise,
but instead sporadic interest. Legislators react to headline
news in both directions—exceptionally horrible crimes on one
side and exceptionally unjust convictions or sentences on the
other. Reexamining the whole system to see that punishments are
actually proportional to crimes is boring work and not on
most legislators’ agenda. This is not a new problem; Blackstone
complained of this legislative “want of attention” two and half
centuries ago.\textsuperscript{102}

A sentencing commission’s advantages of expertise and
sustained interest without the disadvantage (in my view, not
Pfaff’s) of letting legislators off the political hook could be
achieved by outsourcing to the commission the committee work
but not the floor vote. A commission could produce a set of
recommendations, along with suggested amendments to them
by any dissenting commissioners, and the legislature could alter
its rules so that these recommendations are brought directly to
the floor for brief discussion and a prompt vote. Pfaff disagrees
with this kind of approach because letting legislators off the
political hook (i.e., giving voters less say in the criminal law) is
precisely his purpose. If the commission proposals are to come
before the legislature at all, he recommends that they take effect
if the legislature does nothing because this will facilitate “scaling
back punishments” by creating a “less risky” (i.e., less democratic)
path for legislators.\textsuperscript{103}

The idea of delegating legislative authority to bureaucrats who
are not accountable to the people has always been controversial.
The U.S. Supreme Court upheld the constitutionality of delegating
guideline-issuing authority to the Sentencing Commission thirty
years ago, but it did so over a strong dissent by Justice Antonin
Scalia;\textsuperscript{104} his view of delegation has gained many adherents since
that decision. The Supreme Court’s recent \textit{Gundy} decision, in
which four of the eight participating Justices indicated views
along the lines of Justice Scalia’s, may indicate that such broad
delegations are on shaky ground.\textsuperscript{105} Constitutional or not, it
is contrary to the spirit of democracy to have laws that govern
punishment enacted by unaccountable appointees.

Pfaff has a number of other proposals, stated briefly and
mostly in general terms. These include prison closing commissions
similar to the post-Cold War base closing commissions, private
prisons with rehabilitation incentives, sunset clauses on “crime
of the month” laws, the Justice Reinvestment Initiative to fund
alternatives to incarceration, cultural changes in attitudes toward
crime, and social impact bonds.\textsuperscript{106} These are all debatable as to
their cost-effectiveness, but the debate is beyond the scope of this
review given the brevity of their treatment in the book.

The bottom line of the second half of the book is Pfaff’s
proposition that “our attitude toward violent crime needs to change
if we hope to end mass incarceration.”\textsuperscript{107} If weakening both the deterrent and incapacitative effect of our current laws
with a resulting increase in victimization of innocent people
today is the only way to end mass incarceration, then I suspect
that an overwhelming majority of the American people would
just stick with so-called mass incarceration. If there is a path
forward to reducing prison populations, it must not be stained
with innocent blood.

III. The Road Not Taken

But maybe our long-run choice is not between the current
high incarceration rate and inadequately punishing violent
criminals. Perhaps there is another way. There is no Part III to
Pfaff’s book, but there are some tantalizing tidbits that cry out
for further exploration.

In the most curious passage of the book, Pfaff notes that in
the three years ending in 2013 the number of people in prison
for violent crimes fell almost as much as the number in prison for
drug crimes.\textsuperscript{108} Yet, as noted earlier, substantially all “reform”
efforts have been directed to prisoners convicted of nonviolent

\textsuperscript{100} Id. at 216.
\textsuperscript{101} Id. at 218-219.
\textsuperscript{102} 4 William Blackstone, Commentaries 3 (1st ed. 1769).
\textsuperscript{103} Pfaff at 221.
\textsuperscript{104} See Mistretta v. United States, 488 U.S. 361, 413 (1989).
\textsuperscript{105} See Gundy v. United States, 204 L.Ed.2d 522, 537 (2019) (Alito, J.,
concurring in the judgment); \textit{id.} (Gorsuch, J., dissenting).
\textsuperscript{106} See Pfaff 222-228.
\textsuperscript{107} Id. at 229
\textsuperscript{108} Id. at 201.
crimes, with those convicted of violent crimes expressly excluded. “There’s no explanation for how this happened that I’ve seen,” Pfaff says. “[I]n fact, no one really seems to have commented on it at all.”

A plausible explanation fairly screams off the page. The rate of violent crime has fallen dramatically in the last quarter century. The bottom rate of 2014 was less than half of the horrific peak rate of 1991. One would expect a falling crime rate to eventually produce a falling incarceration rate, although it may take considerable time. If an overloaded system was underpunishing violent criminals in 1991, in the view of the people making the punishment decisions, then the fall in incarceration rates would not happen immediately, particularly if the capacity of the system were expanded, as indeed happened. Eventually, though, between capacity expansion and falling crime rates, the system would reach the correct level of punishment; further declines in crime would further decrease the prison population. I do not have the data to test whether this hypothesis is correct, but Pfaff’s statement that no academic is even bothering to ask the question is striking.

If we have reached the point that incarceration rates will fall as crime rates fall, then those whose true goal is reducing the incarceration rate should consider measures that reduce the crime rate as a primary means of achieving the goal. Conversely, if reducing the punishment for violent crime increases the crime rate through diminished deterrence or incapacitation, such measures may be counterproductive to the goal of reducing the incarceration rate.

Pfaff himself recognizes this effect, but curiously only for murder. He says that if police put more emphasis on solving murder cases, then incarceration rates will go up in the short term. He says that if police put more emphasis on solving murder cases, then incarceration rates will go up in the short term. Eventually, though, between capacity expansion and falling crime rates, the system would reach the correct level of punishment; further declines in crime would further decrease the prison population. I do not have the data to test whether this hypothesis is correct, but Pfaff’s statement that no academic is even bothering to ask the question is striking.

If we have reached the point that incarceration rates will fall as crime rates fall, then those whose true goal is reducing the incarceration rate should consider measures that reduce the crime rate as a primary means of achieving the goal. Conversely, if reducing the punishment for violent crime increases the crime rate through diminished deterrence or incapacitation, such measures may be counterproductive to the goal of reducing the incarceration rate.

Pfaff himself recognizes this effect, but curiously only for murder. He says that if police put more emphasis on solving murder cases, then incarceration rates will go up in the short term but down in the long term as more effective enforcement lowers the murder rate, presumably through deterrence, incapacitation, or both. There is no need to limit this to murder. Pfaff points to the recent declines in imprisonment in New York State, noting that the declines came only from New York City, while upstate New York actually increased prison admissions. The decline therefore did not come from statewide legislative changes softening the state’s notoriously harsh drug laws. New York City has seen even more dramatic reductions in crime than the country as a whole, largely because of more effective policing and prosecution.

The intriguing possibility is that prison populations can be brought down through more effective law enforcement without going soft on violent crime. We could invest in more police and more effective policing. More police will, of course, cost a lot of money. More effective use of existing police can be achieved through innovations such as New York’s Compstat and “community policing,” in the original and correct meaning of that term.

While “reformers” often call for funding more law enforcement with the savings from reduced incarceration, it is neither necessary nor practical to expect that we can reduce incarceration before improving law enforcement. Furthermore, savings from reduced incarceration do not come from population reductions alone. It will take closure of facilities and layoffs of personnel to produce real cost savings. It is going to be slow. And if the reductions are achieved through sentence reductions, the resulting increase in crime will offset the gains from better law enforcement to some extent.

Fortunately, it is not necessary to wait for savings, if any, from reducing prisons to fund better law enforcement. The recent improvement of the American economy has state governments flush for now. That money can and should be used for improvements to the institutions that protect people from crime, the first purpose of state and local government. The economy will turn down again eventually. But by the time it does, we should be reaping the benefits of lower crime rates, one of which will be lower imprisonment rates. That cost reduction will help sustain better law enforcement into the future.

Of the factors affecting crime rates, law enforcement is the one most easily changed by government policy, but it is not necessarily the most important. Barry Latzer’s extensive examination of crime rates across time and across groups convincingly demonstrates that culture is at least as important as any other factor in determining crime rates. Pfaff calls for a cultural shift in our attitudes, changing “hearts and minds.” The change he calls for, though, is only in attitudes about punishing crime, not attitudes toward committing crime. A culture of respect for the law and respect for the rights of others would likely do more than any government program to bring down crime rates.

109 Id. at 202.
111 Some people may be talking about reducing the incarceration rate when their real goal is softer punishment of violent criminals for its own sake.
112 Pfaff at 209.
113 Id. at 76.
116 Pfaff at 99.
118 Latzer notes that “[d]efinitions of culture are myriad and often confusing.” He quotes the definition by Geert Hofstede: “the collective programming of the human mind that distinguishes the members of one group or category of people from others.” Id. at 269-270. Perhaps more usefully, he refers to “cultural analyses” in the crime context as studies “which relate the beliefs and values of social groups to their crime rates.” Id. at 265.
120 Pfaff at 227-228.
which would in turn bring down incarceration rates. Yet Pfaff omits any mention of this possibility from his culture change discussion.

IV. Conclusion

The first half of *Locked In* is a very valuable contribution to the field. Pfaff’s thoroughly documented demolition of the Standard Story should be read by everyone concerned with the problems of crime and punishment. Knowing where we really are at present is essential for deciding which direction to go from here. As for our new heading, we can do better than Pfaff’s misguided proposals. The best way to bring down incarceration rates is to bring down crime rates, and more attention is needed on ways to achieve both reductions, not trading one for the other.
I
To understand better the future of the federal judiciary—and why it matters—we should first look to the past. Let’s consider where our judiciary began and how it has gotten to where it is today.

A
I begin where any judge should: with the text, of course. Article III of the United States Constitution established the federal courts and vested in them “the judicial Power of the United States.” But, in contrast to the detail it provides for the powers vested in the other branches, the Constitution’s description of the judicial power effectively stops there. The Constitution identifies those categories of “Cases” and “Controversies” that will be subject to the judicial power of the United States. But it says little about what such power is or how it ought to be exercised.

The concept was not novel to the framers of the Constitution, however. Rather, the general nature of the “judicial power” should have been well known to the founding generation from centuries of experience in England. This included, in the words of Professor Philip Hamburger, the central duty of English judges to “decide [cases] in accord with the law of the land.” That the “judicial Power” was left largely undefined in the new Constitution may simply reflect the fact that its general meaning was already understood.

The traditional conception of the judicial power embodied two related ideals. First, because judges would be deciding cases according to the law, they would not be deciding cases according to their personal values. The law alone was to supply the basis for decision. Legal historians have debated the degree to which this was true in England, disagreeing, for example, over the extent to which English judges would stray beyond the text of a law in the service of more ambiguous principles like equity. But in Federalist 78, Alexander Hamilton defended the proposed Constitution on the very ground that an independent judiciary would help ensure that “nothing would be consulted [in the courts] but the constitution and the laws.”

This critical facet of the judiciary is derived from the unique structure of our government. The Constitution’s structure
importantly tells us what the “judicial Power of the United States” is not: the executive or the legislative powers, which are, of course, vested in the other branches. Unlike in England, where the courts were intertwined with the legislative branch, our judicial power is intentionally separate from the lawmaking powers. Such separation was fundamental to the entire constitutional project. Like Montesquieu before him, James Madison wrote that “no political truth is certainly of greater intrinsic value than [the separation of powers],” and he warned that the failure to divide the legislative, executive, and judicial powers “may justly be pronounced the very definition of tyranny.” As Professor John Manning has persuasively argued, this “reject[i]on of [] English structural assumptions” carries important implications for the nature of our judiciary’s power. Namely, because this structural innovation was made to “limit[] official discretion and promot[e] the rule of law . . . [i]t is difficult to conclude . . . that the Founders also sought to embrace the broad judicial lawmaking powers and discretion” that might have arisen in England.

Second, and related, the Founders envisioned that the “judicial Power” would be exercised in a neutral fashion. Precisely because judges would be, in the words of Hamilton, “bound down by strict rules and precedents, which serve to define and point out their duty,” there would be no “arbitrary discretion in the courts.” Ideally, whether a party prevailed would depend not on the whims of any particular judge, but on the content of the applicable law. In that sense, as Hamilton famously put it, the judiciary was to exercise “neither Force nor Will, but merely judgment.” Again, the Constitution’s structure emphasizes why this must be so. If courts were to behave differently then, in the words of Hamilton, “the consequence would equally be the substitution of their pleasure to that of the legislative body.”

For much of our nation’s history, that traditional ideal remained the dominant conception of judging. But in the 1920s and 30s, scholars began questioning whether achieving the ideal was possible, let alone desirable. Legal realists, as they came to be known, purported to “look[] beyond ideals and appearances for what [was] ‘really going on.’” They argued that judges do not in fact decide cases in accordance with the law because conventional legal materials are too ambiguous or conflicting to yield a single right answer to a case. Thus, the realists argued, judges cannot be trusted when they say the law dictates a particular result; whether judges realize it or not, their decisions rest on considerations outside the law.

Having supposedly debunked the traditional ideal of judging, the realists opened the door to new theories of adjudication. The 1940s and 50s witnessed the rise of the so-called legal process school, which went beyond the realists by developing a new approach to deciding cases in the face of legal indeterminacy. Particularly influential in this respect were the teachings of Professors Henry Hart and Albert Sacks at the Harvard Law School when I was a student there in the 1960s. Hart and Sacks built their theory on the premise that every law has a purpose—a purpose, that is, to address some societal need. It is the essential task of the judge, Hart and Sacks argued, to ensure that these purposes are carried out.

In the latter half of the twentieth century, such approaches to the law grew even more freewheeling and became ascendant. For decades, law schools, the Supreme Court, and the legal profession as a whole were hostile to the traditional view of the judge, which had been replaced by the model of the judge as benign policymaker. This was certainly true in 1963 when I graduated from Harvard during the heyday of the Warren Court, and it was still true in September 1986, when I joined the Ninth Circuit Court of Appeals. Indeed, one of my preeminent colleagues on the court, the late Judge Harry Pregerson, had proudly declared at his confirmation hearing in 1979 that “if a decision in a particular case was required by law or statute . . . offended [his] own conscience,” he would, “try and find a way to follow [his] conscience.” I can assure you that Judge Pregerson did just that, as did many of my colleagues, including the late Stephen Reinhardt who himself advocated for what he described as “judging from the perspective of social justice.” To quote Justice Elena Kagan, at that point in our history, the entire American judicial endeavor was

7 Baron de Montesquieu, The Spirit of Laws, bk. XI, ch. 6, at 163 (Thomas Nugent trans., London, George Bell & Sons 1878) (“[T]here is no liberty if the judiciary power be not separated from the legislative and executive . . . ”).
8 The Federalist No. 47.
9 The Federalist No. 78.
10 Id.
11 Id. (emphasis added).
14 Id. at 196.
15 Id. at 190, 193, 196
17 See Horwitz, supra note 12, at 254 (“The most influential and widely used text in American law schools during the 1950s was The Legal Process by Henry Hart and Albert M. Sacks.”).
19 Id. at 1374.
“policy-oriented” with judges and law students alike “pretending to be congressmen.”

II

In many areas, this freewheeling approach to the law remains alive and well today. I find this troubling, of course, because such views grossly misconceive the “judicial Power” created in Article III and distort the proper role of the judge under our Constitution.

We have all seen the harmful effects of more than a generation of such judicial practice. Any educated American will be familiar with at least some of its more controversial legacies in cases concerning abortion rights, physician-assisted suicide, same-sex marriage, and so on. The chief problem, as I see it, is that these many well-known cases removed social and political questions from the democratic process without a basis in the Constitution’s text or structure. Consider perhaps the most egregious example of social change through judicial fiat: Roe v. Wade, which constitutional scholar and former dean of Stanford Law School John Hart Ely once declared “bad,” not because of its political outcome, but “because it is bad constitutional law, or rather because it is not constitutional law and gives almost no sense of an obligation to try to be.”

Whatever one thinks about the policy outcomes in cases like these, it would be hard to deny the distorting effect they have had on the legitimacy of our federal courts. These cases represent a troubling trend in our country by which litigants, the public, and even members of the bench themselves have come to regard the judicial branch simply as an alternative forum for achieving partisan political goals. We see this starkly in the ugly and personal ways in which we debate judicial nominees today. But perhaps we shouldn’t expect anything different. If courts can deliver results like Roe with hardly a connection to the Constitution, then why wouldn’t these divisive political battles migrate from the Capitol to the courtroom?

Ultimately, I fear, such a trend is unsustainable. It erodes the important divisions between powers erected by our Constitution’s structure, and it raises vital questions about the civic health of our country.

III

I suggest the time has come for a renewed embrace of the Constitution’s limitations on the judicial power to return our courts to their proper role and to reinvigorate our democratic processes. And indeed, there are positive signs on the horizon that change might eventually come.

A

Just as unbound judicial decisionmaking is the primary cause of our current dilemma, a return to sound interpretive principles may be the most promising cure. Let’s consider the two dominant models of restrained judicial interpretation today: textualism and its close relative originalism.

Textualism, at its core, is the simple idea that written statutes should be interpreted according to what their text means. Originalism extends this same idea to constitutional interpretation. In the words of Justice Antonin Scalia, originalists simply believe that the Constitution “means today not what current society, much less the court, thinks it ought to mean, but what it meant when it was adopted.” There is much to be unpacked within these interpretive methods, but for present purposes suffice it to say that each stems from the premise that we should do our best to remove the individual interpreter of a law (i.e., the judge) from the law’s proper interpretation. The meaning of written laws does not depend on the external values of the judge, but instead on the identifiable content of the legal text.

So how do such methods of interpretation reinforce the constitutional design and promote our democratic processes? Primarily by respecting, as opposed to changing, the “legislative bargain”—the deal struck when legislators with competing interests enact law.

Passing law is a messy and haphazard business. To a judge hoping to enforce some lofty purpose behind a legal text, its many idiosyncrasies might seem inexplicable. But a law’s peculiarities are not necessarily its flaws, and textualists enforce the law that the parties actually agreed upon and passed—not the one that some of them, in the court’s view, might have wanted. Enacting public policy requires of legislators a significant commitment of time and other political resources, but textualism promises that the hard-earned fruits of that commitment—i.e., the law itself—will be upheld regardless of the court’s own views on the matter.

This encourages policymakers to do the hard work that their job of governing requires, and it enhances the courts’ legitimacy by keeping them free from partisan political fights. But, more deeply, textualists’ respect for the legislative bargain promotes democratic self-rule and reinforces political accountability. When the textual meaning of a law determines its interpretation, the public knows how and why the law is the way it is, and it knows who can change it: the elected legislators. By contrast, when unelected judges are willing to shape the law themselves, interested parties might find litigation more expedient than engaging in the democratic process. This weakens democratic responsiveness, and it undermines the electoral means by which we normally hold political actors accountable.

Settling these baseline structural questions frees up political actors to focus on addressing the problems of the day and ensures that they will engage in the very democratic mechanisms that exist for them to do so. Professor and former judge Michael McConnell offers a useful analogy: “The rules of basketball do not merely constrain those who wish to play the game, but also make the game possible.”

Speech without grammar is gibberish, and democracy without structure is mob rule. The Founders wrote the rules of the game in 1787, and their rulebook continues to make democratic politics possible in 2019.


B

With this background in mind, let’s focus on the future and see how a return to the traditional role of the judge might one day evolve.

I

The idea that the words of legal instruments determine their meaning is as old as the Constitution itself. But at some point, judges lost their way, and free-ranging judicial methods gained primacy. Textualism and originalism, as theories of interpretation, are still relatively new, developed largely in response to this shift. Indeed, Justice Kagan remarked that, during her time at Harvard Law School in the 1980s, the only interpretive question was typically “what should this statute be,” rather than “what do the words on the paper say.” If someone had mentioned “statutory interpretation” to her while she was in school, she was not sure she “would even quite have known what that meant.”

The good news is that American law has since undergone a sea change. Thanks to the efforts of countless lawyers and judges—and especially to the elevation of Justice Scalia to the Supreme Court—originalism and textualism have become commonplace terms. When Judge Scalia became Justice Scalia, jurists who might once have risked losing their credibility for adopting such legal approaches now had a formidable advocate on our nation’s highest court. He paved the way for judges like me to embrace openly a traditional view of judging that advocated a limited role for the courts. Gradually that view took hold, and the supremacy of an amorphous, outcome-oriented approach to the law waned. Justice Kagan gave perhaps the greatest testament to this seismic shift when she declared in 2015 that, thanks to Justice Scalia, “We’re all textualists now.”

For those of us who remember a time before “textualism” even had a name, this is astounding. Only thirty years after Justice Kagan was taught at Harvard to interpret statutes by what they should say, she described her present approach to the law this way: “When judges confront a statutory text, they’re not the writers of that text; they shouldn’t be able to rewrite that text. There is a text that somebody... has put in front of them, and... what you do with that text is a very different enterprise than the enterprise that Congress... has undertaken in writing that text.”

As I reflect on the state of the federal courts when I first joined the bench, it is remarkable to me that this description of the judicial task is so uncontroversial as to be proudly declared by a Supreme Court Justice—one who was appointed by President Obama, no less!

So that is the first sign that the future of the federal courts might be bright. Textual approaches to the law, though still inconsistently followed, have a real prominence in the legal community today, all the way up to the Supreme Court.

2

The second, and related, sign for hope is in the many promising judges who have recently joined the federal bench.

At the very top, the elevation of Justices Neil Gorsuch and Brett Kavanaugh to the Supreme Court provide for the first time in generations at least five justices with relatively firm commitments to textual interpretations of the law (and, in statutory cases, perhaps more). But the more widespread impact might be in the recent appointments of so many lower court judges who share these same commitments. The combined effect of these new judges could be substantial. Between them, they will decide many thousands of cases—and that means, presumably, many thousands of decisions that will be rooted in a traditional view of the judicial role.

Just as the scores of many contrary decisions shaped the legal community for generations, so too might these decisions have a profound impact on the law and how it is perceived and taught. Even as many law schools remain hostile to these views, I question how long they might continue to produce young lawyers who are not conversant in the language of the law as it is actually written by the judges deciding cases. A concentrated body of textual jurisprudence should clear the way for such methods to be taken seriously in the academy.

Perhaps Stanford Law School is a good example. Here, at the preeminent law school on the west coast and one of the finest schools in the world, you have a Constitutional Law Center directed by one of the leading originalist thinkers in the country: Judge McConnell. The new dean of my alma mater, Harvard Law School, is John Manning, a former law clerk to Justice Scalia and another leading conservative legal thinker. These developments would have been unthinkable when I joined the court. And they have the potential to influence your generation of lawyers and beyond.

And I have no sense that this judicial movement is likely to decline. The thoroughness with which the qualifications, intellectual rigor, and jurisprudential foundations of judicial nominees are reviewed these days is striking. This is not the first time in our history that textualists have been appointed to the federal bench, nor is this the first administration to care about appointing judges of this sort. But, until very recently, these concepts—originalism, textualism, and so forth—simply weren’t the way most people learned, discussed, or thought about the law. They now are, and I believe that can only be a harbinger of good things to come.

IV

In closing, I would like to touch upon something President Reagan said at the investiture of Chief Justice William Rehnquist and Justice Scalia in 1986—the same year I joined the federal bench and the very start of this ongoing judicial movement. There,

26 See, e.g., The Federalist No. 81; Joseph Story, Commentaries on the Constitution §§ 400, 433 (1833).
28 Kagan, supra note 22.
29 Id.
30 Id.
31 Id.
the president mentioned two areas of struggle to nurture and to preserve the structure of our government.

The first struggle is within the judicial branch itself, as judges attempt to stay true to their oaths to “bear true faith and allegiance” to the Constitution, even against the political pressures of the day. Judges must resist the temptation to follow personal preferences over the text of the Constitution—a temptation that is especially great in hard cases, when it’s important to have judges who both care deeply about the Constitution and have the sharpest legal minds. I have been pleased to see that the most recently appointed federal judges seem to have such qualities, and we can hope that those who to continue to join the bench will as well.

The second struggle that President Reagan mentioned is one within the United States at large. At the close of his speech that day, he observed that the entire citizenry must work to preserve the constitutional structure:

We the people are the ultimate defenders of freedom. We the people created the government and gave it its powers. And our love of liberty and our spiritual strength, our dedication to the Constitution, are what, in the end, preserves our great nation and this great hope for all mankind.32

Nurturing this dedication to the Constitution among citizens and within the legal community is a worthy and difficult task. I commend all those who are dedicated to the noble effort of sustaining our founding principles.

---

The Supreme Court Takes Up Abortion: What You Need to Know About June Medical Services v. Gee

By Rachel N. Morrison

Litigation Practice Group

About the Author:
Rachel Morrison serves as Litigation Counsel for Americans United for Life and is a member of the Federalist Society’s Religious Liberties Practice Group Executive Committee. She filed an amicus brief on behalf of Americans United for Life in the case discussed in this article (see note 43).

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the author. Whenever we publish an article that advocates for a particular position, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:

The U.S. Supreme Court’s October 2019 Term started off with a bang. In its first order following the long conference after the Justices’ summer break, the Court agreed to hear cross petitions from a Louisiana abortion provider called June Medical Services and the state of Louisiana stemming from a challenge to Louisiana’s admitting privileges law. The case, June Medical Services v. Gee, raises important issues concerning the future of abortion access and regulations in the United States, the correct application of Whole Woman’s Health v. Hellerstedt, and perhaps even the continuing validity of Roe v. Wade.

I. The Law

The Louisiana law at issue is the Unsafe Abortion Protection Act (or Act 620).¹ The law requires physicians who perform abortions to have active admitting privileges at a hospital within 30 miles of the facilities where they perform abortions.² A physician has “active admitting privileges” if he or she “is a member in good standing of the medical staff” of a licensed hospital, “with the ability to admit a patient and to provide diagnostic and surgical services to such patient.”³

The purpose of the law, as discussed throughout the state’s briefing, is threefold. First, it creates uniformity in the law by bringing abortion providers under the same requirements that already applied to physicians providing similar types of services at other ambulatory surgical centers. Second, the law performs a credentialing function. Since hospitals perform more rigorous and intensive background checks than do abortion clinics in Louisiana, requiring a physician to have admitting privileges at a hospital ensures that the physician has the requisite skills and capacity to perform relevant procedures—in this case, abortions. Third, the law helps ensure that women who suffer complications from abortion procedures receive continuity of care by enabling the direct and efficient transfer of both the patient and her medical records to a local hospital.

II. The Lawsuit

On August 22, 2014—after Louisiana passed the law and prior to its effective date of September 14, 2014—June Medical Services, along with two other Louisiana abortion clinics and two Louisiana abortion doctors,⁴ filed a lawsuit in the Middle District of Louisiana requesting that the law be enjoined because it allegedly placed an undue burden on their patients’ access to


³ Id.

⁴ During the course of the litigation, the two other abortion clinics shut down (for reasons unrelated to Act 620) and dropped out of the case. For ease of reference, I refer to all plaintiffs as “June Medical.”
abortion. In Planned Parenthood of Southeastern Pennsylvania v. Casey, the Supreme Court established the undue burden standard to determine whether an abortion regulation violates the Constitution. “[A]n undue burden is a shorthand for the conclusion that a state regulation has the purpose or effect of placing a substantial obstacle in the path of a woman seeking an abortion of a nonviable fetus.”6

The district court entered a temporary restraining order, keeping Louisiana’s abortion facilities from going into effect during preliminary injunction proceedings.7 After a bench trial, the court granted a preliminary injunction, holding that the admitting privileges requirement was facially unconstitutional and enjoining enforcement of the law.8 Louisiana’s request to the district court to stay the injunction pending appeal was denied, but its request to the Fifth Circuit Court of Appeals for an emergency stay pending appeal was granted.9 The Fifth Circuit explained, “Louisiana is likely to prevail in its argument that [June Medical] failed to establish an undue burden on women seeking abortions or that the Act creates a substantial obstacle in the path of a large fraction of women seeking an abortion.”10 The court also noted that a pending Supreme Court case—Whole Woman’s Health v. Hellerstedt11—involved a nearly identical admitting privileges law in Texas.12 The following day, June Medical filed an application in the Supreme Court to vacate the Fifth Circuit’s stay.13 A week later, the Supreme Court granted June Medical’s application and vacated the Fifth Circuit’s stay.14

III. The Supreme Court’s Intervening Hellerstedt Decision

At the end of June 2016, the Supreme Court issued its decision in Hellerstedt. By a 5-3 vote (Justice Antonin Scalia passed away shortly before the opinion came down), the Court invalidated two provisions of Texas’ H.B. 2, which required abortion doctors to have admitting privileges at a local hospital and abortion facilities to follow certain surgical-center standards.15 These provisions were unconstitutional, the Court said, because they created an undue burden on abortion access.16

Notably, the Hellerstedt Court modified Casey’s undue burden standard by requiring that “courts consider the burdens a law imposes on abortion access together with the benefits those laws confer.”17 After weighing the benefits and burdens of Texas’ law, the Court ultimately invalidated the two provisions because “[e]ach place[d] a substantial obstacle in the path of women seeking a previability abortion.”18 Citing the record 22 times, the majority opinion explained that the district court “applied the correct legal standard” when it “considered the evidence in the record—including expert evidence, presented in stipulations, depositions, and testimony.”19

After Hellerstedt came down, the Fifth Circuit remanded the case back to the district court to “engage in additional fact finding required by” Hellerstedt.20 On April 25, 2017, the district court entered final judgment and permanently enjoined the law.21 After weighing the evidence, the district court “found that Act 620 confers only minimal” health benefits, but “substantial burdens,” and ruled that, on its face, “Act 620 places an unconstitutional undue burden on women seeking abortion in Louisiana.”22

IV. Fifth Circuit Decision

On appeal, the Fifth Circuit reversed and ruled 2-1 in favor of Louisiana’s law, explaining that there were “stark differences” between the facts and evidence in the Texas case and the facts and evidence in Louisiana’s case.23 Unlike in Texas, there was no evidence that any abortion clinic would close in Louisiana as the result of the law.24 After a detailed examination of the factual record, the Fifth Circuit concluded that Act 620 would—at worst—cause up to one hour of delay for abortion procedures at one of Louisiana’s three clinics.25

June Medical appealed to the en banc Fifth Circuit, but the judges voted 9-6 to deny rehearing the case en banc.26 The court also denied a stay pending appeal.27

---

6 Id. at 877.
10 Id. at 328.
11 136 S. Ct. 2292 (2016).
12 See June Med. Servs., 814 F.3d at 328 n.16 (noting that the interests at issue in Hellerstedt were not implicated in the case).
15 136 S. Ct. at 2300.
16 Id.
17 Id. at 2309; see also id. at 2310 (stating that the district court applied the correct legal standard when it “weighed the asserted benefits against the burdens”).
18 Id. at 2300.
19 Id. at 2310.
22 Id. at 86.
24 Id.
25 Id. at 813.
V. EMERGENCY STAY PENDING APPEAL

The same day the Fifth Circuit denied the stay request—January 25, 2019—June Medical made an emergency stay request to the U.S. Supreme Court, asking the Court to stop Louisiana's law that was set to go into effect on February 4 from being enforced while a petition for certiorari was submitted to the Court.28

In order for the Supreme Court to put Louisiana's law on hold while the case was being appealed, there had to be: (1) a “reasonable probability” that the Court (i.e., four Justices) would agree to take the case; (2) a “fair prospect” that a majority of the Justices would ultimately find the law unconstitutional; and (3) a “likelihood of irreparable harm” that would result if the stay was denied.29

Louisiana opposed June Medical’s stay request, arguing that the law should not be put on hold because this is not the type of case the Court will normally agree to take since June Medical did not identify any conflict in the circuit courts and its disagreement with the Fifth Circuit panel is mainly over how best to interpret the facts.30

The request was made to Justice Samuel Alito as the Justice in charge of emergency requests from the Fifth Circuit, and he referred it to the full Court. On February 1, Justice Alito ordered an “administrative stay,” or a temporary hold, through Thursday, February 7 on Louisiana’s law to give the Justices more time to review the arguments made by June Medical and Louisiana.31 The order specified that this temporary hold “does not reflect any view regarding the merits” of the case.32

Late Thursday night, just hours before Louisiana's admitting privileges law would go into effect, the Court agreed 5-4 to grant June Medical's emergency stay request, putting Louisiana’s law on hold while the case is appealed.33

No rationale was given for the Court’s decision, which is normal for emergency requests. And despite dissenting in *Hellerstedt*, Chief Justice John Roberts joined the Court's more liberal justices—Justices Ruth Bader Ginsburg, Stephen Breyer, Sonia Sotomayor, and Elena Kagan—in (presumably) agreeing that there was a “reasonable probability” the Court would agree to take the case and ultimately find the law unconstitutional.

Justices Clarence Thomas, Samuel Alito, Neil Gorsuch, and Brett Kavanaugh would have denied the abortion providers’ request and allowed Louisiana’s law to go into effect. Justice Kavanaugh wrote a dissent, pointing out the many “factual uncertainties” involved in the case and saying there was no reason at that time for the Court to stay the law because if abortion doctors in Louisiana really could not obtain admitting privileges, they could file as-applied challenges at that point.34

According to the order, the stay on Louisiana's law would automatically be lifted if the case was not timely appealed, if the Court decided not to take the case after all, or if the Court issued a final judgment.35

VI. CERT PETITION AND CONDITIONAL CROSS-PETITION

In April 2019, June Medical filed a petition for certiorari.36

The question presented was: “Whether the Fifth Circuit's decision upholding Louisiana's law requiring physicians who perform abortions to have admitting privileges at a local hospital conflicts with this Court's binding precedent in *Hellerstedt*.”37

June Medical argued that the decision below conflicts with the Supreme Court’s decision in *Hellerstedt*, which struck down a nearly identical admitting privileges law as unconstitutional.38

They claimed that Louisiana’s law lacks health and safety benefits and will burden women seeking abortions in Louisiana. Therefore, under *Hellerstedt*’s requirement to “consider the burdens a law imposes on abortion access together with the benefits those laws confer,” the “non-existent benefits are outweighed by its extensive burdens.”39 June Medical even went so far as to tell the Court that summary reversal is appropriate and that the Fifth Circuit disregarded binding precedent.40

Louisiana opposed June Medical’s petition, arguing that the Fifth Circuit made no legal error and emphasizing the multiple complex issues of fact and law, which made the case procedurally unsuited to further review.41 If the Court did grant review, Louisiana said it would only be appropriate to clarify or limit *Hellerstedt*.42

8A774/81802/20190125210126@26-2%20Order%20Opposing%20Motion%20for%20Stay.pdf.


32 Id.


34 Id. at 4 (Kavanaugh, J., dissenting).

35 Id. at 1 (majority opinion).


37 Id. at i.

38 Id.

39 Id. at 31 (first quotation quoting *Hellerstedt*, 136 S. Ct. at 2309).

40 Id. at 32–35.


42 Id. at 36–39.
In addition to opposing June Medical’s petition, Louisiana filed a conditional cross-petition, arguing that if the Court agrees to take the case, it should also consider whether abortion providers can be assumed to have third-party standing to challenge health and safety regulations, such as Louisiana’s admitting privileges law.43

Ordinarily, parties must bring a lawsuit on their own behalf, but sometimes third parties can bring a lawsuit on behalf of another. Usually, the Court’s third-party standing doctrine requires: (1) a “close” relationship between the third party and the person who possess the right, and (2) a “hindrance” to the possessor’s ability to protect his own interests.”44 But this changed in the abortion context after the Supreme Court’s decision in Singleton v. Wulff, in which the Court stated that “it generally is appropriate to allow a physician to assert the rights of women patients as against governmental interference with the abortion decision.”45 Since then and based on this generality, many lower courts and even the Supreme Court have generally assumed that abortion providers have third-party standing on behalf of women seeking abortions without any meaningful, particularized analysis (as is required in other contexts) of whether there is a close relationship between abortion providers and their patients and a hindrance to the patients’ ability to sue on their own behalf.46

Louisiana also raised the issue in its conditional cross-petition of whether objections to prudential standing (including third-party standing) are waivable or not, pointing to a circuit split.47

June Medical opposed the cross-petition, arguing that Louisiana had waived its challenge to third-party standing, that third-party standing is subject to waiver, and that there is no underlying circuit split for the court to resolve.48 They argued that settled precedent establishes that abortion providers have third-party standing and there is no reason for the Court to revisit the issue.49

VII. COURT GRANTS CERT ON BOTH PETITIONS

On October 4, 2019—the first day orders were issued from the Justices’ long conference after the summer break—the Court granted both petitions for certiorari and consolidated the cases for briefing and one hour of oral argument.50 The questions presented are:

1. Whether abortion providers can be presumed to have third-party standing to challenge health and safety regulations on behalf of their patients absent a “close” relationship with their patients and a “hindrance” to their patients’ ability to sue on their own behalf;

2. Whether objections to prudential standing are waivable; and

3. Whether the U.S. Court of Appeals for the 5th Circuit’s decision upholding Louisiana’s law requiring physicians who perform abortions to have admitting privileges at a local hospital conflicts with the Supreme Court’s binding precedent in Whole Woman’s Health v. Hellerstedt (2016).

Oral argument will likely be set for late winter or early spring 2020.

VIII. ISN’T LOUISIANA’S LAW THE SAME AS TEXAS’ LAW IN HELLERSTEDT?

The first thing usually mentioned about this case is that Louisiana’s law is materially similar or identical to the Texas law that the Supreme Court found unconstitutional in 2016 in Whole Woman’s Health v. Hellerstedt.51 But the Court’s ruling in Hellerstedt does not mean that all admitting privileges laws are per se unconstitutional or that there is sufficient evidence in the record that Louisiana’s law will lead to the closure of a large number of abortion clinics in Louisiana. Determining whether an abortion regulation is unconstitutional under the undue burden test is a fact-intensive inquiry that requires state-specific evidence that the law causes a substantial obstacle to abortion access. Therefore,


Full disclosure: I filed an amicus brief on behalf of Americans United for Life (AUL), where I serve as Litigation Counsel, arguing that abortion providers should not be assumed to have third-party standing to bring legal challenges against health and safety violations on behalf of their patients. See Brief Amicus Curiae of Americans United for Life in Support of Cross Petitioner, Gee v. June Med. Servs. L.L.C., No. 18-1460 (Vide 18-1323), (U.S. June 24, 2019), https://aul.org/wp-content/uploads/2019/06/18-1460-Amicus-Brief-of-Americans-United-for-Life.pdf. AUL’s brief explains that June Medical brought the current legal challenge against a backdrop of serious health and safety violations by Louisiana abortion clinics and professional disciplinary actions against and substandard medical care by Louisiana abortion doctors. The violations and disciplinary actions by Louisiana abortion providers documented in the brief demonstrate that June Medical does not have a close relationship with their patients and should not have third-party standing:

There is an inherent conflict of interest between abortion providers and their patients when it comes to state health and safety regulations. It is impossible for abortion clinics and doctors to share or represent the interests of their patients when they seek to eliminate the very regulations designed to protect their patients’ health and safety. Id. at 3–4.


46 Cf. Hellerstedt, 136 S. Ct. at 2322 (Thomas, J., dissenting) (“A plurality of this Court fashioned a blanket rule allowing third-party standing in abortion cases.”).

47 Conditional Cross-Petition at i.


49 Id.


51 136 S. Ct. 2292.
the Justices will look at all of the specific factual nuances in the record to determine whether this case is *Hellerstedt* 2.0 or if there are “stark differences” between Texas and Louisiana, as the Fifth Circuit held.

IX. Who Has Standing?

Regarding the first question presented, the assumption of third-party standing for abortion providers has been called into question by academics and judges alike, including most notably Justice Thomas. In Thomas’ *Hellerstedt* dissent, he stated:

> The Court’s third-party standing jurisprudence is no model of clarity. Driving this doctrinal confusion, the Court has shown a particular willingness to undercut restrictions on third-party standing when the right to abortion is at stake. And this case reveals a deeper flaw in straying from our normal rules: when the wrong party litigates a case, we end up resolving disputes that make for bad law.

Given comments like this, Justice Thomas may jump at the opportunity to provide clarity to the Court’s third-party standing doctrine in the abortion context.

If the Court clarifies its doctrine on standing and requires that there be a close relationship between abortion providers and their patients and a hinderance to their patients’ ability to sue on their own behalf in order for abortion providers to legally challenge an abortion regulation, it will presumably kick the case back down to the district court to decide in the first instance whether June Medical has standing to challenge Louisiana’s law. The case could then be resolved by the Supreme Court on a procedural issue clarifying the standard for third-party standing, without an actual determination on the merits of the constitutionality of Louisiana’s law, or even whether June Medical does or does not have standing in this particular case.

If, however, the Court decides that abortion providers can be presumed to have third-party standing to challenge abortion regulations on behalf of their patients or that objections to prudential standing are waivable, such that it is too late for Louisiana to raise a challenge to June Medical’s standing, it would presumably reach the merits on the third question presented—whether the Fifth Circuit’s decision to uphold Louisiana’s admitting privileges law conflicts with *Hellerstedt*.

X. WHAT IS THE CORRECT INTERPRETATION OF *Hellerstedt*?

The third question presented in the case would allow the Court to clarify the correct interpretation and application of *Hellerstedt*. Since the Court’s decision in 2016, lower courts and parties have disagreed over what *Hellerstedt* requires when it says that a court must consider “the burdens a law imposes on abortion access together with the benefits those laws confer.”

Pro-abortion groups urge a broad reading, claiming that if there is no (or a de minimis) benefit of the law, any demonstrated burden—no matter how small—renders the law unconstitutional. Several pro-abortion groups have also brought a novel challenge under *Hellerstedt*, arguing that a state’s entire abortion regulatory scheme, or a group of a state’s abortion laws, cumulatively create an undue burden. This new claim is referred to as a “cumulative burden claim” or “cumulative effects challenge.”

On the other hand, states defending their abortion regulations urge a more narrow reading of *Hellerstedt*, pointing out that the Court explicitly relied on *Casey* when it invalidated Texas’ law and that *Casey’s* standard “asks courts to consider whether any burden imposed on abortion access is ‘undue.’” Thus, a regulation on abortion cannot be unconstitutional unless the law places a substantial obstacle in the path of a woman seeking an abortion and its “numerous burdens substantially outweigh[] its benefits.”

*Hellerstedt* has created confusion for state legislators who are unsure what type of abortion-related health and safety laws (if any) they can pass. If the Court gets to the merits or at least opines on what the standard of review is for determining the constitutionality of abortion regulations, *Hellerstedt*’s requirements should be made clearer to parties, judges, and state legislators.

Four of the five Justices in the *Hellerstedt* majority are still on the Court: Justices Ginsburg, Breyer (the author), Sotomayor, and Kagan. The three dissenting Justices remain as well: Chief Justice Roberts and Justices Thomas and Alito. There are two new Justices: Justices Gorsuch replaced Justice Scalia, who passed away shortly before the *Hellerstedt* opinion was issued, and Justice Kavanaugh replaced Justice Kennedy, who joined the majority. This case presents the first opportunity for both Justices Gorsuch and Kavanaugh to rule on the merits of an abortion decision addressing the application of *Hellerstedt*, *Casey*, and *Roe*.

None of the four Justices in the *Hellerstedt* majority will likely disagree with that opinion, especially considering they voted otherwise.
to grant June Medical’s emergency stay, but the new opinion could provide more clarity as to what *Hellerstedt* requires. It is, however, an open question whether the four Justices will be able to obtain a fifth vote. Interestingly, Chief Justice Roberts voted to grant the emergency stay of Louisiana’s law pending appeal to the Court. It is unclear whether he did this because he has reconsidered his earlier dissent in *Hellerstedt* or for some other reason.

XI. Conclusion

Court-watchers are paying attention. For many, how the Court chooses to resolve this case, including its interpretation and application of *Hellerstedt*, will indicate the direction the Court is moving on the abortion issue.
In *Carpenter v. United States*, the United States Supreme Court confronted an issue at the crossroads of technology, societal notions of privacy, and the meaning of the Fourth Amendment. Its resolution of that issue brought into stark relief profound disagreements among the Justices concerning constitutional construction, the nature of judicial precedent, and indeed the meaning of judging itself. Since the Supreme Court decided *Carpenter* in 2018, a number of reviewing courts—state and federal—have considered its myriad potential implications. They have not yet scratched the surface, and *Carpenter* stands today both as a conceptual challenge for practitioners and judges, and quite possibly a landmark of Fourth Amendment jurisprudence.

And it all starts, mundanely enough, with a string of electronics store robberies.

I. The *Carpenter* Decision

In 2011, four men were arrested in Detroit for robbing several Radio Shack and T-Mobile stores in the area. Investigators soon learned that the robberies were not limited to Detroit. Indeed, a ‘suspect identified 15 accomplices who had participated in the heists and gave the FBI some of their cell phone numbers; the FBI then reviewed his call records to identify additional numbers that he had called around the time of the robberies.’

Prosecutors sought court orders for the cell phone records of Timothy Carpenter and others pursuant to the Stored Communications Act. The Act allows the government access via compulsory process to particular telecommunications records maintained by private entities, so long as the government can show, to the satisfaction of a federal magistrate, “specific and articulable facts showing that there are reasonable grounds to believe that . . . the records . . . are relevant and material to an ongoing investigation.” Specifically, prosecutors sought to compel disclosure of cell-site data from MetroPCS and Sprint:

Those data themselves took the form of business records created and maintained by the defendants’ wireless carriers: when the defendants made or received calls with their cellphones, the phones sent a signal to the nearest cell-tower for the duration of the call; the providers then made records, for billing and other business purposes, showing

---

2 *Id.* at 2212.
3 *Id.*
4 *Id.*
5 *Id.*
which towers each defendant’s phone had signaled during each call.\(^7\)

The orders were applied for and allowed, and thus “the government obtained 12,898 location points cataloging Carpenter’s movements—an average of 101 data points per day.”\(^8\)

Carpenter was subsequently charged with six counts of robbery and an assortment of firearm offenses.\(^9\) Carpenter moved before trial to suppress the data provided by the carriers, arguing that their seizure violated the Fourth Amendment where it was obtained without a warrant, and the district court denied the motion.\(^10\) Carpenter was convicted on all counts save one firearm motion.\(^10\) Carpenter was convicted on all counts save one firearm count, and the Sixth Circuit affirmed his conviction in a published opinion, holding, among other things, that Carpenter, according to well-established United States Supreme Court precedent, had no expectation of privacy in cell phone records created, stored, and maintained by a third party.\(^11\) The Supreme Court granted Carpenter’s petition for certiorari.\(^12\)

In an opinion authored by Chief Justice John Roberts, the Supreme Court reversed the Sixth Circuit.\(^13\) The Court began by noting that, contrary to earlier precedent, modern Fourth Amendment jurisprudence is not mechanically tethered to pure questions of property law and the common law doctrine of trespass, that is, actual physical intrusions by the government onto the property of another.\(^14\) Instead, the Court has established that the Fourth Amendment protects people, not places, and expanded [its] conception of the Amendment to protect certain expectations of privacy as well. When an individual seeks to preserve something as private and his expectation of privacy is one that society is prepared to recognize as reasonable, [the Court] has held that official intrusion into that private sphere generally qualifies as a search and requires a warrant supported by probable cause.\(^15\)

Indeed, the majority analyzed the case with head-on reference to this well-settled but (as we shall see) much-criticized “reasonable expectation of privacy” test.\(^16\) The Court proceeded to observe that the kind of data at issue—historical cell-site location information, or CSLI, maintained by a third party—“does not fit neatly under existing precedents. Instead, requests for cell-site records lie at the intersection of two lines of cases, both of which inform our understanding of the privacy interests at stake.”\(^17\)

The first line of cases, the Court noted, concerns a person’s expectation of privacy in his physical location and movements. In United States v. Knotts, for example, the Court held in 1983 that police use of a “beeper” tracking device secretly placed by them in a container and later acquired by Knotts and unknowingly placed by him in his own vehicle violated no reasonable expectation of privacy.\(^18\) The Court in Knotts made the commonsense observation that someone “traveling in an automobile on public thoroughfares has no reasonable expectation of privacy in his movements from one place to another” and that, because those movements had been “voluntarily conveyed to anyone who wanted to look,” there simply was no “search” in the constitutional sense.\(^19\) Knotts was distinguished and refined in 2012, after decades of technological progress and the advent of more sophisticated law enforcement tools and techniques. In United States v. Jones, the Supreme Court held that Fourth Amendment protections applied where federal law enforcement secretly installed a GPS tracking device on Jones’ Jeep Grand Cherokee and monitored its location and movements for 28 days.\(^20\) The Court in Jones straightforwardly held that the unconsented-to surreptitious attachment of the GPS device to Jones’ personal property—his Jeep—was an actual physical occupation of private property by the government in an effort to acquire information and was thus a search within the meaning of the Fourth Amendment, the only issue before the Court.\(^21\) Nonetheless, five Justices went on to argue in dicta that, setting aside the actual physical trespass by the government, the GPS tracking of Jones implicated his constitutional privacy interests as contemplated by Katz.\(^22\) Furthermore, “[s]ince GPS monitoring of a vehicle tracks every movement a person makes in that vehicle, the concurring Justices concluded that longer term GPS monitoring in investigations of most offenses impinges on expectations of privacy—regardless of whether those movements were disclosed to the public at large.”\(^23\)

The second line of cases, the Carpenter Court observed, deals with what has become known as the “third-party doctrine.”\(^24\) That doctrine stands for the proposition that “a person has no legitimate expectation of privacy in information he voluntarily turns over to third parties.”\(^25\) Thus, information such as records of phone numbers dialed from a person’s home\(^26\) or a person’s banking

\(^7\) United States v. Carpenter, 819 F.3d 880, 885-886 (2016), reh’g en banc denied, June 29, 2016.

\(^8\) Id.

\(^9\) Id.

\(^10\) Id.


\(^12\) Carpenter, 138 S. Ct. at 2223.

\(^13\) Id. at 2213.

\(^14\) Id. (internal quotations and citations omitted).

\(^15\) That test was originally articulated by the Court in Katz v. United States, 389 U.S. 347, 351 (1967), and it has been applied by courts construing the Fourth Amendment ever since.

\(^16\) Carpenter, 138 S. Ct. at 2214-15.


\(^18\) Id. at 281.


\(^20\) Id. at 404-05.

\(^21\) Carpenter, 138 S. Ct. at 2215 (citing Jones, 565 U.S. at 426, 428).

\(^22\) Id. (quoting Jones, 565 U.S. at 430 (internal quotations and citations omitted)).

\(^23\) Id. at 2216.

\(^24\) Id.


\(^26\) Id.
records”—the subjects of Smith and Miller, respectively—have traditionally received no Fourth Amendment protection whatsoever. This was a bright line rule that, to many practitioners and courts, had the oft-sought virtue of being relatively simple to apply, even if its faithful application sometimes led to counterintuitive results. This was sometimes the case because, so long as the information was voluntarily disclosed to a third party, the Constitution was not implicated, “even if the information [wa]s revealed on the assumption that it [would] be used only for a limited purpose.”

In Carpenter, then, the Court was faced with the question of how to treat CSLI in the light of both strands of cases. The second strand presented what could be considered a threshold question: Where law enforcement can track an individual’s past movements by scrutinizing a record of his cell phone signals created and maintained by his wireless carrier, does a straightforward application of the third-party doctrine necessitate an equally straightforward result of no Fourth Amendment protection? The Court said no:

Given the unique nature of cell phone location records, the fact that the information is held by a third party does not by itself overcome the user’s claim to Fourth Amendment protection. Whether the Government employs its own surveillance technology as in Jones or leverages the technology of a wireless carrier, we hold that an individual maintains a legitimate expectation of privacy in the record of his physical movements as captured through CSLI.

The Court went on to explain—tying in the first strand of cases—that a person’s expectations of privacy are not surrendered simply because she conducts her affairs and moves about in public. Citing Katz and its reasonable expectation of privacy standard, the Court observed that “what [one] seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected.” Recalling the GPS tracking in Jones and the concerns expressed by that case’s concurrence, the Court reiterated that tracking a person’s public movements for an extended period of time intrudes on a reasonable expectation of privacy, even if that tracking takes the form of business records created and maintained by a third-party commercial entity, such as a wireless provider:

Although such records are generated for commercial purposes, that distinction does not negate Carpenter’s anticipation of privacy in his physical location. Mapping a cell phone’s location over the course of 127 days provides an all-encompassing record of the holder’s whereabouts. As with GPS information, the time stamped data provides an intimate window into a person’s life, revealing not only his particular movements, but through them his familial, political, professional, religious, and sexual associations.

The Court proceeded to expound on the fact that new technology allows for more sweeping surveillance than was considered in its prior cases. As law enforcement capabilities grow, the sphere of protection provided to a person’s reasonable expectations of privacy must grow commensurately. The more the government can do, the more the Constitution must do to keep pace. “With access to CSLI,” the Court argued:

the Government can now travel back in time to retrace a person’s whereabouts subject only to the retention policies of the wireless carriers, which currently maintain records for up to five years. Critically, because location information is continually logged for all of the 400 million [wireless] devices in the United States—not just those belonging to persons who might happen to come under investigation—this newfound tracking capacity runs against everyone. Unlike with the GPS device in Jones, police need not even know in advance whether they want to follow a particular individual, or when.

Moreover, the Court noted that it is inaccurate in this context to say that a person voluntarily and knowingly discloses his location information to his third-party provider simply by carrying a cell phone. The Court observed that cell phones are ubiquitous in modern life, and that an active cell phone generates its own location information without the need for any affirmative action by its holder. Indeed:

Virtually any activity on the phone generates CSLI, including incoming calls, texts, or e-mails and countless other data connections that a phone automatically makes when checking for news, weather, or social media updates. Apart from disconnecting the phone from the network, there is no way to avoid leaving behind a trail of location data.

The Court then held that it would not “extend” the third-party doctrine as set forth in Smith and Miller to the collection of CSLI, finding that CSLI is sui generis and its gathering by the government, a search. Moreover, where the acquisition of CSLI is a search, that search must be authorized by a warrant supported by probable cause. The Court made sure to declare that its holding was a “narrow one,” and that it was expressing no views on issues not expressly before it in Carpenter. The Court also made clear

---

28 Id. at 443.
29 Carpenter, 138 S. Ct. at 2217.
30 Id. (quoting Katz, 389 U.S. at 351-352).
31 Id. (internal citation and quotations omitted).
32 Id. at 2219 (“There is a world of difference between the limited types of personal information addressed in Smith and Miller and the exhaustive chronicle of location information casually collected by wireless carriers today.”). See also Orin S. Kerr, An Equilibrium-Adjustment Theory of the Fourth Amendment, 125 Harv. L. Rev. 476 (2011) (setting out a theory of how the Supreme Court continually modifies and refines the scope of Fourth Amendment protections in response to social and technological developments), available at https://harvardlawreview.org/2011/12/an-equilibrium-adjustment-theory-of-the-fourth-amendment/.
33 Carpenter, 138 S. Ct. at 2218.
34 Id. at 2220.
35 Id.
36 Id. at 2221.
37 Id. at 2220.
that its holding did “not disturb the application of Smith and Miller or call into question conventional surveillance techniques and tools, . . . .” 38 Nonetheless, the majority opinion found itself faced with a panoply of dissents from four Justices, raising issues of the most fundamental and contentious sort.

II. The Carpenter Dissents

Justice Anthony Kennedy initiated the gauntlet of dissents with an opinion joined by Justices Clarence Thomas and Samuel Alito. Kennedy’s opinion emphasized that, properly understood, Carpenter was simply about the government’s use of congressionally authorized compulsory process to obtain relevant business records in the usual course of a criminal investigation. 39 Process was allowed, pursuant to the Stored Communications Act, by a neutral and detached magistrate, after the government demonstrated that the records were reasonably necessary to an ongoing investigation. 40 Yet the majority had determined that this was not a simple demand for records from a third party, but a search in the constitutional sense affecting the rights of a person who was plainly not the holder of the documents subject to compelled disclosure. This, to Kennedy and the Justices who joined him, was unprecedented:

In concluding that the Government engaged in a search, the Court unhinges Fourth Amendment doctrine from the property-based concepts that have long grounded the analytic framework that pertains in these cases. In doing so it draws an unprincipled and unworkable line between cell-site records on the one hand and financial and telephonic records on the other. 41

According to Kennedy, there is no way to make a distinction, in any constitutionally cognizable way, between someone’s credit card records and their CSLI. Both open to investigators a window into a person’s life that he has already revealed to the record keeper. And under Smith and Miller, that revelation should be dispositive in Carpenter. Carpenter’s CSLI records are pure business records, and Carpenter “could expect that a third party—the cell phone service provider—could use the information it collected, stored, and classified as its own for a variety of business and commercial purposes.” 42 Carpenter had no property interest in the company’s records, and to say that he nonetheless maintained a privacy interest in them makes no sense and departs from well-settled Fourth Amendment jurisprudence. The bright line has been muddied, if not erased, and what had been a straightforward analytical framework was demolished by the wrecking ball of the allegedly “entirely different species of business record” that is CSLI.

Justice Alito wrote his own dissent, joined by Justice Thomas, and was even more critical. Justice Alito noted that the majority’s decision elided the important distinction between actual

physical searches—where government agents enter and search, say, someone’s home or office—supported by probable cause, and an order to produce records:

Treating an order to produce like an actual search, as today’s decision does, is revolutionary. It violates both the original understanding of the Fourth Amendment and more than a century of Supreme Court precedent. Unless it is somehow restricted to the particular situation in the present case, the Court’s move will cause upheaval. Must every grand jury subpoena duces tecum be supported by probable cause? 43

Certainly such a proposition would work a sea change for law enforcement, but in his rigorously argued dissent, Justice Alito leaves the reader wondering how such an outcome does not follow ineluctably from the majority’s reasoning and premises.

Alito proceeded with a comprehensive historical tour of the role of compulsory process in American law from the time of the founding. He demonstrated that its use was never considered a search, and that probable cause was never required for its issuance. Simply put, compulsory process did not historically fall within the ambit of the Fourth Amendment. That amendment instead—and according to its own words—simply prohibits unreasonable searches of an individual’s “person, house, papers, and effects.” Thus, Fourth Amendment law traditionally incorporated a property-based component consistent with common law notions of trespass. 44 So by its terms,” Alito concluded:

the Fourth Amendment does not apply to the compulsory production of documents, a practice that involves neither any physical intrusion into private space nor any taking of property by agents of the state. Even Justice Brandeis—a stalwart proponent of construing the Fourth Amendment liberally—acknowledged that “under any ordinary construction of language,” “there is no ‘search’ or ‘seizure’ when a defendant is required to produce a document in the orderly process of a court’s procedure.” 45

The showing necessary for a compelled production of documents, as Justice Kennedy observed, is a straightforward one of relevance and reasonableness, not the probable cause required for search warrants. 46 There is no question here, according to Alito, that the order for Carpenter’s CSLI, authorized by the Stored Communications Act, fell comfortably within the constitutional strictures for compulsory process. 47

Finally, Alito delivered a devastating and apparently unanswerable critique of the majority:

Against centuries of precedent and practice, all that the Court can muster is the observation that “this Court has never held that the Government may subpoena third parties for records in which the suspect has a reasonable expectation

---

38 Id.
39 Id. at 2224.
40 Id.
41 Id.
42 Id. at 2230.
43 Id. at 2247.
44 Id. at 2247.
45 See Jones, 565 U.S. at 405.
46 Carpenter, 138 S. Ct. at 2251 (quoting Olmstead v. United States, 277 U.S. 438, 476 (1928) (dissenting opinion)).
47 Carpenter, 138 S. Ct. at 2255.
of privacy.” Frankly, I cannot imagine a concession more damning to the Court’s argument than that. As the Court well knows, the reason that we have never seen such a case is because—until today—defendants categorically had no “reasonable expectation of privacy” and no property interest in records belonging to third parties.48

Thus the circular logic of the majority on this crucial analytical point comes into clear and, as Justice Alito aptly puts it, damning, relief. Moreover, Alito went on to explain how the majority misapprehends Miller and Smith and what has become known as the third-party doctrine. He noted that the third-party doctrine was never a new doctrine at all, but was merely a consistent and logical application of Fourth Amendment first principles. The idea that one can object to a governmental intrusion upon the property of another flies in the face of the Fourth Amendment’s history and language, where persons are protected in their persons, houses, papers, and effects.

Justice Thomas penned a remarkable solo dissent, in which he questioned why the Court uses Katz at all: “The Katz test has no basis in the text or history of the Fourth Amendment. And, it invites courts to make judgments about policy, not law. Until we confront the problems with this test, Katz will continue to distort Fourth Amendment jurisprudence.”49 Justice Thomas began by recounting the Katz test’s unlikely evolution from almost an impromptu afterthought at oral argument in 1967, through Justice Harlan’s concurrence where the phrase “expectation of privacy” first appears in American jurisprudence, to its full-throated adoption by the Court as the “lodestar” for determining whether a constitutional search occurred in Smith.50 He proceeded to explain why Katz’s holding that a search occurs whenever the government violates someone’s reasonable expectation of privacy cannot be squared with the text and original meaning of the Fourth Amendment. A search at the time of the founding had a different meaning: an actual, physical search of a home or office or other location by agents of the government. Moreover, the text of the Fourth Amendment specifically protects people in their persons, houses, papers, and effects, not simply any place a person may have a reasonable expectation of privacy. And individuals have a right to be secure in their persons, houses, papers, and effects, not those of others. Finally, to leave it to a court to decide whether someone’s expectation of privacy is reasonable is simply asking for trouble in terms of clarity, predictability, and other legal values.

As Justice Scalia famously observed, “In my view, the only thing the past three decades have established about the Katz test . . . is that, unsurprisingly, those actual (subjective) expectations of privacy that society is prepared to recognize as reasonable bear an uncanny resemblance to those expectations of privacy that this Court considers reasonable.”51 Finally, Justice Thomas urged the Court to abandon the Katz test wholesale.52 The majority opinion noted, however, that no party in Carpenter asked the Court to revisit Katz.53

Justice Gorsuch, in his own erudite dissent, criticized both Katz and the third-party doctrine, advocating for a more traditional, property law-based approach to the Fourth Amendment. In his view, courts deciding cases like this should look not to their own opinions or preferences, but to accepted positive-law sources such as statutes.54

III. Applying Carpenter

Since Carpenter was released in the summer of 2018, several reviewing courts and various trial courts have grappled with its implications, though none have yet crossed the minefields telegraphed by the dissents. In United States v. Hood, the defendant, charged with the transportation and receipt of child pornography, argued that the Internet Protocol (IP) address information that the government obtained from the smartphone messaging company Kik without a warrant should be suppressed under Carpenter.55 The First Circuit disagreed and held that, unlike the CSLI in Carpenter, IP address information by itself conveys no information about a person’s location:

The IP address data is merely a string of numbers associated with a device that had, at one time, accessed a wireless network. By contrast, CSLI itself reveals—without any independent investigation—the (at least approximate) location of the cell phone user who generates that data simply by possessing the phone a cell phone.

In contrast, an internet user like the defendant in Hood makes the “affirmative decision to access a website or application.”56 This distinction was enough for the First Circuit to find Carpenter inapplicable, and the court noted its agreement with the only other circuit court to have addressed the issue post-Carpenter.57

Several federal district court opinions have taken the law and logic of Carpenter in directions more amenable to defendant

52 Carpenter, 138 S. Ct. at 2246. As evidenced by some recent opinions, Justice Thomas is not shy about urging that the Court reconsider some venerable cases and doctrines that he believes, and attempts to demonstrate, rest on particularly unstable foundations. See, e.g., McKee v. Cosby, 139 S. Ct. 675 (Thomas, J., concurring in denial of cert.) (Feb. 19, 2019) (explaining reasons for reconsidering the constitutional requirement that public figures satisfy an actual-malice standard for state-law defamation actions); Garza v. Idaho, 139 S. Ct. 738, 756-59 (Thomas, J., dissenting) (questioning the underpinnings of the constitutional right to effective assistance of counsel at taxpayer expense).
53 Carpenter, 138 S. Ct. at 2214, n.1.
54 Id. at 2267-72.
55 United States v. Hood, 920 F.3d 87 (1st Cir. 2019).
56 Id. at 92.
57 See United States v. Contreras, 905 F.3d 853, 857 (5th Cir. 2018) (“The information at issue here falls comfortably within the scope of the third-party doctrine. [The] records revealed only that the IP address was associated with the Contreras’s Brownwood residence. They had
the new context of historical CSLI. And where historical CSLI and historical GPS are functionally equivalent for purposes of a Fourth Amendment analysis, binding appellate precedent did not authorize the warrantless acquisition of the GPS data at issue. The district court finally concluded that, although the car contract explicitly stated that the dealer was authorized to use an embedded GPS tracking device to track the car's whereabouts, as with Carpenter's cell phone, it could not be said that the users of the car truly “voluntarily turned over” that data to any third party.

Several district courts have also dealt with stationary surveillance cameras post-Carpenter. In United States v. Kelly, the Eastern District of Wisconsin held that stationary video surveillance of the exterior of an apartment building and the hallway outside of an apartment for forty-nine days did not require a warrant under Carpenter. The court noted:

> Unlike a cell phone, the video surveillance did not track the totality of the defendant’s movements. It tracked only his arrival to and departure from a residence that wasn’t his. The defendant’s attempt to equate a process that records only what someone standing in the apartment hallway, or outside the apartment complex, could have seen with a process that follows a person into homes, places of worship, hotels, bedrooms, restaurants and meetings, takes Carpenter’s reasoning too far.  

In contrast, the District of Massachusetts recently held in United States v. Moore-Bush, which involved a stationary surveillance camera attached to an outside utility pole, that a defendant “had an objectively reasonable expectation of privacy in their and their guests’ activities around the front of [their] house for a continuous eight-month period.” The court said:

> It stands to reason that the public at the time of the [Fourth Amendment’s] framing would have understood the King's constables to violate their understanding of privacy if they discovered that constables had managed to collect a detailed log of when a home's occupants were inside and when visitors arrived and whom they were.

Although the surveillance in Moore-Bush was considerably longer than that in Kelly, and although the camera in Moore-Bush was trained on the defendant's home as opposed to someone else's or a common hallway, both cases implicate a standard investigative technique whose lawful limits are now called into question by Carpenter and its doctrinal ancestors. The various courts of appeals will have to grapple with these issues soon, no doubt in anticipation of further refinement and explication by the Supreme Court.

Finally, the state court whose 2014 opinion presaged Carpenter in most material respects confronted an issue the Supreme Court specifically did not address in Carpenter: real-time CSLI. In Commonwealth v. Almonor, the Supreme Judicial Court of Massachusetts held that the government’s causing a defendant’s cell phone to reveal its real-time location by “pinging” the phone—that is, having the service provider cause the phone to transmit its GPS coordinates to the provider—is a search under Article 14 of the Massachusetts Declaration of Rights (the state’s equivalent of the federal Fourth Amendment). A warrant supported by probable cause was therefore required. Although the court held that exigent circumstances excused the failure to obtain a warrant in Almonor, it is now clear that such an investigative technique is a search in Massachusetts.

The reasoning of the majority opinion in Carpenter provided valuable jurisprudential support for the state court’s holding.

59 Id. at 653-54.
60 See Davis v. United States, 564 U.S. 229, 249-50 (2011) (holding that “when the police conduct a search in objectively reasonable reliance on binding appellate precedent, the exclusionary rule does not apply”).
61 Diggs, 385 F. Supp. 3d at 657.
62 Id.
63 Id. at 660.
64 Id. at 660.  
66 Id. at 727.
68 Id. at 148.
69 See Commonwealth v. Augustine, 467 Mass. 230, 251 (2014) (holding that government acquisition of CSLI must be by a search warrant supported by probable cause because defendant had a reasonable expectation of privacy in the CSLI under Article 14 of the Massachusetts Declaration of Rights).
70 482 Mass. 35 (2019).
71 Id. at 52-53.
IV. Conclusion

Lower courts have barely scratched the surface of *Carpenter* and its implications. The many opinions in the case are a feast of passionate argument, legal philosophy, and American history. One thing is for certain: the rules of *Carpenter* and *Katz* will lend fuel to the fires of legal debate in courtrooms and the academy for a long time to come.
In a collection of speeches, judicial opinions, and anecdotes, Justice Neil Gorsuch's new book offers advice to legal and lay audiences alike on the importance of civility, courage, and humility while weaving in his views on the separation of powers, originalism, and textualism, among other legal issues. He draws inspiration for these lessons in life and law from former bosses including Justices Byron White and Anthony Kennedy, legal heroes such as Justices John Marshall Harlan and Robert Jackson, and his family, law clerks, friends, and many colleagues. The book also offers a glimpse into the private world of a man who was catapulted from relative obscurity in Colorado to the national stage with his appointment to the Supreme Court.

For Justice Gorsuch, civility is a cornerstone of our republic. Without it, “the bonds of friendship in our communities dissolve, tolerance dissipates, and the pressure to impose order and uniformity through public and private coercion mounts.”

Self-governance “turns on our treating each other as equals—as persons, with the courtesy and respect each person deserves—even when we vigorously disagree.” It's a quality his former boss Justice Kennedy instilled in him (“one can disagree but never disagreeably”), and he, in turn, hopes to instill it in his law clerks. He saw it in action during his recent confirmation to the Supreme Court, and he shares stories of the many acts of kindness he experienced—a care package with socks and a note that his looked worn out on television, a joke told while he was in line getting coffee, and well wishes from someone across the political aisle. They are proof that “goodness . . . runs deep in our collective history and sustains our republic.”

Justice Gorsuch also highlights the courage many great American lawyers and judges have shown and their willingness “to stand firm for justice in the face of immense pressure and often at grave personal costs.” He points to John Adams' willingness to represent British soldiers following the Boston Massacre in 1770, Justice Harlan's lone dissent in *Plessy v. Ferguson* (writing that the Constitution “neither knows nor tolerates classes among citizens”), and Justice Jackson's dissent from the Court's rulings in the *Chenery* cases. In each instance, these men knew their actions could alienate friends and harm their reputations. Adams, Harlan, and Jackson are all models of courage for Justice Gorsuch and proof that adhering to the law “in the face of great public pressure is sometimes a lonely business.” That lonely road is one worth walking, however, and judges should aspire to be humble in carrying out their duties.

---

2 Id.
3 Id. at 312.
4 Id. at 182.
5 Id. at 24.
It’s apparent from the very cover of the book that Justice Gorsuch takes humility seriously. He lists as collaborators his former law clerks Jane Nitze and David Feder, who worked in his chambers at the Tenth Circuit as well as the Supreme Court. They helped the Justice sort through countless speeches and judicial opinions to select a sampling of Gorsuch’s greatest hits. This recognition of their contributions is not something one would expect to see in a book authored by a Supreme Court Justice. On the point of humility, Justice Gorsuch shares a story of walking through the halls of the Supreme Court with his boss Justice White. Justice White admitted he only knew about half of his predecessors as they passed their portraits. “We’ll all be forgotten soon enough,” Justice White told him, and Gorsuch concludes that “this is exactly as it should be . . . most any of us who believe in [our republic’s] cause can hope for it that we have done, each in our own small ways, what we could in its service.”

An outgrowth of this humility is Justice Gorsuch’s firm belief that judges must avoid the temptation to rule for certain groups or policy outcomes. When judges rule according to their personal preferences rather than the law, he notes, “[t]he people are excluded from the lawmaking process, replaced by a handful of unelected judges who are unresponsive to electoral will, unrepresentative of the country . . .” In his view, “[v]irtually the entire anticanon of constitutional law we look back upon today with regret came about when judges chose to follow their own impulses rather than follow the Constitution’s original meaning.”

The opinion excerpts included in the book showcase how Justice Gorsuch practices what he preaches—disagreeing with colleagues without being disagreeable, staking out positions that may be unpopular, and advocating for a judiciary that stays within its limits. In an excerpt from Henson v. Santander Consumer USA, his maiden majority opinion for the Supreme Court, Justice Gorsuch considers a situation where the Court was asked to act like a legislature. The case involved whether a loan purchaser can be considered a debt collector under the Fair Debt Collection Practices Act, which by its text regulates debt collection agencies and not the loan originators who hired them. “Faced with obstacles in the text and structure,” Justice Gorsuch writes, “petitioners ask us to move quickly on to policy.” They pressed the Court to “update” the law passed in the 1970s given changes in the industry in the intervening decades. Declining that invitation, the Justice explains, “we will not presume with petitioners that any result consistent with their account of the statute’s overarching goal must be the law” and instead “presume more modestly” that the legislature says “what it means and means . . . what it says.” It is, after all, “never our job to rewrite a constitutionally valid statutory text under the banner of speculation about what Congress might have done.”

In United States v. Nichols, Gorsuch dissented from a Tenth Circuit ruling upholding the federal Sex Offender Registration and Notification Act that delegated to the Attorney General the power to determine how, when, and whether the law’s registration requirement would apply to sex offenders convicted before the law went into effect. Gorsuch writes forcefully that this is a clear violation of the separation of powers: “[i]f the separation of powers means anything, it must mean that the prosecutor isn’t allowed to define the crimes he gets to enforce.”

Though it has been “some time” since the Supreme Court ruled that a law “cross[ed] the line” from a permissible delegation to a violation of the nondelegation doctrine (more than 80 years, in fact), it “has also been some time,” then-Judge Gorsuch notes, “since the courts have encountered a statute like this one.” Upholding this law would require the Judiciary to endorse the notion that Congress may effectively pass off to the prosecutor the job of defining the very crime he is responsible for enforcing. By any plausible measure we might apply that is a delegation run riot, a result inimical to the people’s liberty and our constitutional design.

In Gutierrez-Brizuela v. Lynch, then-Judge Gorsuch wrote a concurring opinion asserting that the Chevron and Brand X deference doctrines violate the separation of powers. These doctrines “permit executive bureaucracies to swallow huge amounts of core judicial and legislative power and concentrate federal power in a way that seems more than a little difficult to square with the Constitution of the framers’ design.” This is “a problem for the judiciary” as well as “the people whose liberties may now be impaired not by an independent decisionmaker . . . but by an avowedly politicized administrative agent seeking to pursue whatever policy whim may rule the day.”

A number of the opinion excerpts demonstrate Justice Gorsuch’s view that the judiciary should play a limited, but important, role in our government. For Justice Gorsuch, following the original meaning of laws or constitutional provisions “is the very reason we have independent judges: not to favor certain groups or guarantee particular outcomes, but to ensure that all persons enjoy the benefit of equal treatment under existing law as adopted by the people and their representatives.” He dispels the notion that this approach “inevitably” leads to rulings in favor of preordained political outcomes. “Rubbish,” he writes, “Originalism is a theory focused on process, not on substance. It is not ‘Conservative’ with a big C focused on politics. It is conservative in the small c sense that it seeks to conserve the meaning of the Constitution as it was written.” The text itself is “the natural starting point for resolving any dispute over its

6 Id. at 16.
7 Id. at 44.
8 Id. at 115.
9 Id. at 221.
10 Id. at 222.
11 Id.
12 Id. at 87.
13 Id. at 95.
14 Id.
15 Id. at 76.
16 Id. at 79.
17 Id. at 10.
18 Id. at 114-15.
He shows originalism and textualism in action in several excerpted opinions. In a dissent from Carpenter v. United States, the Justice discusses the original meaning of the Fourth Amendment and the asexual “reasonable expectation of privacy” standard developed in Katz v. United States. He writes, “[f]rom the founding until the 1960s, the right to assert a Fourth Amendment claim didn’t depend on your ability to appeal to a judge’s personal sensibilities about the ‘reasonableness’ of your expectations of privacy. It was tied to the law.” The Fourth Amendment safeguards “the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures.” “True to those words and their original understanding,” Gorsuch notes, “the traditional approach asked if a house, paper or effect was yours under law. No more was needed to trigger the Fourth Amendment.” This protection “do[es] not depend on the breach of some abstract ‘expectation of privacy’ whose contours are left to the judicial imagination . . . [it] grants you the right to invoke its guarantees whenever one of your protected things . . . is unreasonably searched or seized. Period.” He laments that many litigants forfeit these arguments, “leaving courts to the usual hand-waving.”

In United States v. Rentz, writing for the en banc Tenth Circuit, then-Judge Gorsuch considers whether 18 U.S.C. § 924(c) allows multiple charges stemming from one act. The statute mandates five years’ imprisonment for “any person who, during and in relation to any crime of violence or drug trafficking crime . . . uses or carries a firearm, or who, in furtherance of any such crime, possesses a firearm . . . .” He explains, “in the statute’s language we find three relevant verbs: uses, carries, and possesses. This alone supplies some evidence that each charge must involve an independent act of using, carrying, or possessing.” Reading the statute “in accord with the normal uses of statutory (and sentence) construction goes some way to suggest that every new conviction requires a new act falling into one of those three categories.” He reasons:

Just as you can’t throw more touchdowns during the fourth quarter than the total number of times you have thrown a touchdown, you cannot use a firearm during and in relation to crimes of violence more than the total number of times you have used a firearm . . . . Unless and until [Congress amends the statute], we will not relegate men and women to prison . . . because they did something that might—or might not—have amounted to a violation of the law as enacted.

Justice Gorsuch’s lessons on civility, courage, and humility are as relevant for laymen as they are for law students, lawyers, and judges. This sampling of his most important judicial opinions offers insight into how the Justice puts his commitment to originalism, textualism, and judicial restraint into practice. Beyond that, Justice Gorsuch offers a rare glimpse into his private world. He shares poignant stories about his final moments of anonymity before being thrust onto the national stage with his nomination to the Supreme Court. He reveals how a neighbor helped him evade reporters camped out near his home in the Colorado countryside and how he enjoyed coloring pictures with a little girl on the plane ride to Washington that would change his life forever. He writes, “Yes, I had written hundreds of judicial decisions over the last decade, sitting on an appellate court that serves about 20 percent of the continental United States. But few people outside of legal circles knew who I was. That life was now over.” He also shares stories about tagging along to work with his mother (“a feminist before feminism”) who was the first female lawyer in the Denver District Attorney’s Office, how his father imparted his love of the outdoors (especially fishing), and how his British wife came to love life in the Great American West. There’s also a selection of photographs of his family, their many pets, and his happiest memories (fishing with his daughters). While fans of Justice Gorsuch will enjoy reading his speeches and opinions, it’s the brief piercing of the judicial veil that leaves the reader wanting more. This Gorsuch fan hopes a sequel is in the works.
National security concerns cut across typical arguments about trade policy. Typical arguments concentrate on effects on domestic businesses, workers, and consumers from changes in import flows—flows that are determined by differences between national pay scales, regulatory regimes, and saving and consumption priorities. In contrast, national security concerns look to specific effects on national capacity to protect against perils with potential for broad, national impact.

One increasingly discussed security focus today concerns imports from China, notably imports of certain equipment in the information and communication technology (ICT) sector. News commentary has been mostly confined to one or two firms’ products. Yet the combination of China’s political system, economic structure, and export orientation poses broader threats, which are magnified in a world in which national security is inextricably connected to ICT, both because of its significance to a range of economic, financial, and ordinary daily functions and because of its integration with military hardware and operations. While some aspects of this threat may be controlled through decisions on government procurement and investment, other aspects require constraints on imports of a broader set of products. In addition to products from Huawei and ZTE (the Chinese firms most frequently discussed in connection with security risks), products of other commercially successful firms in the ICT sector—such as China Mobile, Lenovo, and Lexmark—would be on the list.1

This paper reviews the risks posed by Chinese imports, the conditions leading to these risks, the firms and products that could pose these risks, and ways in which those risks might best be addressed, particularly through invocation of Section 232 of the Trade Expansion Act of 1962.

I. Trade Basics: Economics and Security

Trade expands choices, enhances competition, raises standards of living, and increases wealth for both trading parties.2 Anyone who recalls a time when people were limited to fruits and vegetables that were grown nearby appreciates the benefits of access to a worldwide market that can ship products from warmer climates to markets that are experiencing winter, and anyone who grows these products can appreciate having consumers in other markets eager to buy their goods. Having the broadest possible choice set of products that suit each person’s tastes, interests, and budget is easy to like. That is why trade’s overall effects are strongly positive, even though it undeniably can have negative

---

1 See infra, section III.

2 See, e.g., Milton Friedman, Capitalism and Freedom 13 (1962) (making the general case for exchanges between willing participants).

For explanations of the fundamentals of international trade, including the principal driving forces behind trade flows, see, e.g., Jagdish N. Bhagwati & T.N. Srinivasan, Lectures on International Trade chaps. 1–8 (1983).
effects on some businesses, workers, and communities. Over the past seventy-plus years, reductions in global trade barriers, largely associated with the General Agreement on Tariffs and Trade (GATT), have helped expand global trade at roughly half again the pace of global GDP and contributed to major increases in income and declines in poverty.3

Most arguments for trade restrictions rooted in concerns about economic dislocation elevate transient, concentrated effects associated with any change in economic factors—primarily changes in costs of production, technology, and consumer tastes—above broader, longer-term gains to society; such arguments have played out in different terms over more than two centuries. But careful scholars, including those known as proponents of managed trade in specific settings, recognize the strong, general case for open trade and reasons for caution in restricting it.4

One set of concerns, however, is different and has been recognized as a special ground for setting aside normal trade rules. Article XXI of the GATT (brought forward into the World Trade Organization (WTO) framework) provides:

> Nothing in this Agreement shall be construed . . . . (b) to prevent any [member country] from taking any action which it considers necessary for the protection of its essential security interests . . . (iii) taken in time of war or other emergency in international relations . . .

The precise meaning of the GATT language is debated, especially the degree to which the italicized phrase precludes WTO dispute resolution bodies from second-guessing a member state’s judgment of its security needs. But the point of the provision is clearly to mark out a special limitation on interference with a nation’s protection of its security, including self-protection through otherwise prohibited restrictions on trade.

II. RISKS TO U.S. NATIONAL SECURITY

Communications among government personnel engaged with national security issues always have been sensitive, high-priority targets for infiltration by actual and potential opponents. They have also been high-priority for protection through encryption and other steps to reduce opportunity for interception, translation, and defensive or retaliatory maneuvers. For example, breaking the German military’s codes used in its Enigma machines often is credited as contributing significantly to Allied forces’ success in defeating Nazi Germany in World War II.6

In today’s world, communications are even more important and far more numerous and constant. Their importance is partly tied to the vast increase in use of electronic transactions—including, but not limited to, in the domain of finance—in place of what formerly required physical operations. Further, much of what still takes place in the physical realm—such as driving a car or a tank, piloting a plane, or sending missiles toward targets—is governed by instructions that are communicated at a distance or by processes taking place within physically separate equipment pursuant to integrated circuits’ memory and computing processes.7

Any process that incorporates computer chips and any process that occurs at the direction of an electronically transmitted instruction is potentially vulnerable to cyber-espionage and cyber-warfare.8 In the age of the internet, that covers virtually all of our important, our everyday, and our highly sensitive functions. Diplomatic, strategic, and tactical communications and operations necessary to national security are vulnerable to concentrated hacking efforts, potential sources of leakage of communications, and possible weaknesses in the internal instruction sets that govern computing functions.9 Every nation’s protection depends on the robustness of the insulation around these electronic operations.

All of us are familiar with weaknesses in the way that data are collected, stored, and transmitted. When one of our credit cards is hijacked, it could be because of a major leak of data from a company we’ve done business with, or a thief could have stolen the information necessary to access our accounts from a single transaction at a terminal in a store. Even though we are notified that our data may have been taken and cancel the card, we are left to wonder when the theft occurred and what damage may have been done that will not surface right away. Our nation’s secret communications and the security of critical equipment may be subject to even greater risks, as the resources trained on intercepting or disrupting those functions may be far greater and far more focused than those deployed in the commercial realm.

In addition to the risks to national security from efforts to exploit weaknesses in government equipment, software, and communications, serious security risks attach to equipment, software, and communications of government contractors and others with whom the government does business. The risks include not only those associated with direct efforts to exploit weaknesses in communicating and computing, but also those from latent or even unknown weaknesses in communicating and computing resources that interface with government directly or as links in a larger chain. A “backdoor” may be built into commercially successful software or embedded in equipment that is widely

---


8 See, e.g., id. (This message was emphasized throughout the GAO report, explaining GAO’s decision to subtitle its report “DOT and Industry Have Efforts Under Way, But DOT Needs to Define Its Role in Responding to a Real-World Attack.”).

available in ordinary consumer markets, allowing access to highly sensitive information stored on computing or communications equipment directly or remotely, or possibly providing a key to opening other connections leading to such information. Given the number of governments and other entities around the world with interest in discovering information held by the United States or in restricting U.S. military, diplomatic, or other operational options, it is entirely appropriate for the U.S. government to adopt a highly protective stance toward reducing these risks.

III. CHINA TRADE'S SECURITY RISKS

Many nations and many entities pose risks. China and Chinese-origin products, however, pose special risks because a combination of several factors increases the possibility of the products' use for purposes harmful to U.S. security.

The first factor is China's announced goal of dominance in numerous fields, including ICT, that are critical to security, intra-government communications, and military effectiveness. China has made no secret of its intentions in this respect and has made extensive investments in support of these goals.

Second, China has made broad and intense investments in espionage, both in China and abroad, notably including cyber-espionage. It has extensive networks of espionage assets, human and technical, deployed in China and increasingly overseas. This underlies cautions issued by the U.S. government to officials and business executives traveling in China and having on-going communications with Chinese citizens.

Third, China's economy, although still evolving, is not driven by large numbers of small, independent, privately-run firms. Instead, unlike most of the major world economies, it depends on a very large degree on state-owned enterprises (SOEs) and firms that, while not formally state-owned, rely for funding on major (often controlling) investments from the Chinese government. There are estimates to be more than 150,000 SOEs in China, including in some of China's largest enterprises, apart from government investments in many if not most nominally private enterprises that are engaged in substantial economic activity. These enterprises often are led by former government functionaries, including high-ranking members of China's communist party and former military officers. While these officials may no longer have direct roles in government, there is at the least reasonable suspicion of their continuing ties to and responsiveness to the government.

Fourth, also in contrast to most successful and almost all advanced national economies, China's political regime is both openly authoritarian and insulated against formal democratic checks on its exercise of government power. Although for at least a quarter-century China loosened controls over various economic decisions and activities, China's government under President Xi has been reasserting control over many aspects of China's economic activity. As one observer reported, “Since 2012, private, market-driven growth has given way to a resurgence in the role of the state.” The reassertion of control over the economic sector has gone hand-in-hand with assertion of greater control over other activity, including renewed restraints on public speech.

10 See, e.g., Bob Flores, The Dangers of Backdoor Software Vulnerability and How to Mitigate Them, Cyber Defense (May 7, 2019), available at https://www.cyberdefensemagazine.com/the-dangers-of-backdoor-software-vulnerabilities-and-how-to-mitigate-them/ (observing that “as the complexity and scale of application development has advanced, and the components and dependencies have expanded . . . the attack surface [for backdoors] is significantly broader” and the decreasing cost of computing and storage dramatically facilitate cyber-attack options).


18 See, e.g., Maidland & Chatzky, supra note 17 (observing that the "government has considerable sway over all Chinese private companies" because of heavy regulation and government-connected executive appointments). See also Wendy Lutten, Firm Control: Governing the State-Owned Economy Under Xi Jinping, 2018 CHINA PERSPECTIVE 27 (issue 1-2, 2018) (exploring the relationship between consolidation of personal power and greater state control over economic activity).


and publicly available information. Recent events in Hong Kong are merely the most widely observed evidence of these changes.

Part of the framework in place in China under the current regime is the legal requirement that private firms cooperate in government security initiatives, including by granting access to private communications and fully cooperating with China’s Cyberspace Administration. This creates special risk for anyone using telecommunications, computing, or related equipment from a broad array of well-known Chinese firms including Huawei, ZTE, China Mobile, Lenovo, and Lexmark, among others. All of these firms have considerable investment from or control by China’s government, leadership that is intimately connected to China’s government or military, and evidence of product or service features that raise specific questions regarding intended or coincidental security risks.

A final factor in the riskiness of Chinese ICT imports is that these firms’ products typically are complex, sophisticated, and technologically advanced—characteristics that increase opportunities for inclusion of features that can be exploited with or without the firms’ active cooperation. The risks posed by such products are considerably greater, and less easily evaluated, than risks associated with ordinary commercial purchases of less complex products, such as the glass used to make mobile phone screens. Even such highly sophisticated products can pose relatively low security risks, as it is much more difficult to manipulate features to permit state espionage or related intrusions.

It is important to recognize the possibility that any of the above factors could be overstated due to a lack of sufficient credible information. Overstatement also can occur because personal interests may be served by exaggeration of risks or manipulation of factual information. With respect to risks associated with ICT products from China, however, there is at least as great a prospect that the risks are understated rather than overstated. There are obvious interests for the government of China, entities associated with the government, firms that produce and export ICT products from China to the United States, and entities that currently sell or use such products (or wish to) to minimize any estimation of the security risks associated with commercially viable and often low-priced China-sourced products.

Attention to error rates and error costs is essential to critical analysis, and caution before taking a complaint about imports as gospel is sensible. Yet the nature and importance of national security risks, the manifest connection of complex ICT products to such risks, and the complex of factors that make China-sourced ICT products especially likely to pose such risks together provide strong basis for setting aside the usual reservations about pleas for limiting imports or for regulating their use.

IV. POTENTIAL REMEDIES

There are several possible remedies to the risks posed by China-sourced ICT products. While not an exhaustive listing, some of the major candidates are described below.

One obvious remedy is to make changes to U.S. government procurement rules to guard against inclusion of such products in departments and operations of special sensitivity. But such changes are unlikely to be availing. Security lapses often have been traced to government officials’ personal equipment—not to their work-purchased equipment and services—or to the equipment and networks of non-government personnel (particularly government contractors). These lapses can be addressed by strengthening enforcement of rules respecting government personnel’s use of equipment or services even for strictly personal communications. But highly publicized lapses in security by officials at the highest levels—lapses that occurred despite security personnel’s cautions about the activities that led to them—suggest the difficulty of reliance on such rules. Moreover, there simply are too many points of interaction between government and non-government actors in respect of even very sensitive security-related functions to gain much traction through limits on government purchasing and government personnel alone.

Another potential remedy that addresses part of the problem just noted is to amend rules governing government contractors.


Some have already called for such changes. See, e.g., Beeny, et al., supra note 24; Keiser & Smith, supra note 24, at 15, 26; U.S. Department of Defense Inspector General, supra note 13.
as well as government personnel.\footnote{29 See, e.g., Beeny, et al., supra note 24, at 33; Keiser & Smith, supra note 24, at 15.} Here, too, some gains may be had, but the same enforcement problems that attach to attempts to enforce security regulation through rules addressed to government officials stand in the way of effective control of security risks through regulations aimed at government contractors. Asking a broad swath of entities and individuals who work for private firms that do business with the U.S. government to appreciate the risks from use of widely available commercial ICT products is apt to be insufficient protection of national security. The breaches of security that have been traced back to officials’ privately owned products, to equipment and services of government contractors, and to the personnel who work for those contractors are sufficiently numerous to highlight the difficulty of directing others what products to use and how to assure their security.

A different and broader possible remedy would rely on imposing restrictions on importation and sale in the United States of certain China-sourced ICT products that are deemed to pose significant risks to the security of the United States. The most likely vehicle for effecting such restrictions would be Section 232 of the Trade Act of 1962.\footnote{19 U.S.C. § 1862.} The provision, as amended, requires the Secretary of Commerce, in consultation with the Secretary of Defense and other appropriate government officials, to conduct an investigation of the possible national security effects of particular imports (when requested by particular parties or on his own initiative). The Secretary evaluates the effects of the imports on national security and recommends to the President whether and what action is appropriate to eliminate or reduce adverse security effects. The President is given broad discretion to determine whether the imports threaten U.S. national security. He also is granted expansive authority to determine the appropriate action if he decides those imports do threaten national security, including negotiating limits on imports but also extending to an unspecified wider range of options.

On its face, Section 232 seems to offer a clear option for the U.S. to investigate ICT imports and their impact on U.S. security interests and, if necessary, to address the threats through import restraints or other means. While the U.S. law’s plain text would cover actions restricting importation and sale of ICT products that could compromise U.S. security by virtue of their potential susceptibility to espionage from China, some arguments about the law reach back to the underlying international trade provision that determines whether Section 232’s implementation would be consistent with U.S. obligations under the GATT and WTO.\footnote{31 See, e.g., Brandon J. Murrill, The “National Security Exception” and the World Trade Organization, \textit{Cong. Research Serv.} (Nov. 28, 2018), available at https://fas.org/sgp/crs/row/LSB10223.pdf.} The provision at issue states that the GATT does not prevent any member country “from taking any action which it considers necessary for the protection of its essential security interests.”\footnote{GATT art. XXI, sec. b, supra note 5.} That provision is followed directly by three clauses listing reasons a nation might conclude that its interests are threatened. The third clause covers actions deemed necessary to protect security that are “taken in time of war or other emergency in international relations.”\footnote{34 See, e.g., Russia — Measures Concerning Traffic in Transit (DS512), Responses of the United States of America to Questions from the Panel and Russia to Third Parties (GATT Dispute Resolution Proceeding), Feb. 20, 2018, at 1–5, available at https://ustr.gov/sites/default/files/enforcement/DS/US.3d.Prvs.Pnl.Am.Rus.Chn.fin.%28public%29.pdf.} The United States takes the position that what a nation “considers necessary” for the protection of its essential security interests is up to each nation, as the phrase’s emphasis not on what is necessary but on what a nation \textit{considers} necessary strongly suggests.\footnote{35 See, e.g., William A. Reinsch, \textit{The WTO’s First Ruling on National Security: What Does It Mean for the United States?}, \textit{Center for Strategic \& Int’l Studies} (Apr. 5, 2019), available at https://www.csis.org/analysis/wios-first-ruling-national-security-what-does-it-mean-united-states.} A recent decision of a WTO dispute resolution panel rejected that reading, but there is considerable doubt whether that particular ruling will be upheld.\footnote{36 See, e.g., Keiser \& Smith, supra note 24; Maillard \& Chatzky, supra note 17; Five Eyes Will Not Use Huawei in Sensitive Networks, Reuters, Apr. 24, 2019, available at https://www.reuters.com/article/us-britain-huawei-ncsc-usa-five-eyes-will-not-use-huawei-in-sensitive-networks-senior-uk-official-idUSKCN1S01CZ; Czech Cyber Watchdog Calls Huawei, ZTE Products a Security Threat, Reuters, Dec. 17, 2018, available at https://www.reuters.com/article/us-czech-huawei/czech-cyber-watchdog-calls-huawei-zte-products-a-security-threat-idUSKBN1OG1Z3.} Moreover, even if the WTO decides that it is authorized to decide the necessity of actions to respond to an “emergency in international relations,” there certainly is a strong argument that national security threats tied to escalating cyber-espionage and prospects for cyber-espionage satisfy Article XXI’s conditions.

If the U.S. initiates a proceeding under Section 232, finds a national security threat, and undertakes actions designed to restrict imports of Chinese ICT products that might present security risks, political pushback from China is almost inevitable. Chinese officials have been vocally opposed to restrictions on products from Huawei and ZTE which have been identified by several nations, including the United States, as conducive to Chinese cyber-espionage.\footnote{Huawei Products a Security Threat, Reuters, Dec. 17, 2018, available at https://www.reuters.com/article/us-czech-huawei/czech-cyber-watchdog-calls-huawei-zte-products-a-security-threat-idUSKBN1OG1Z3.} If limitations are imposed on a wider array of items from a larger group of firms, the level of complaints from China certainly would rise. In response, China would likely impose sanctions against U.S.-sourced exports to China and increase efforts to persuade U.S. firms dependent on China trade to vocally oppose the government’s actions. Given China’s recent willingness to wield its economic muscle and its political control of the law and markets within China to secure favorable results, there is substantial reason to expect some U.S. firms to voice support for China’s position in any trade conflict.\footnote{Apart from the self-interest of firms seeking to advance their own prospects of favorable treatment in China, there is ample reason to expect China to use its economic clout outside China as a source of advantage. See, e.g., Center for Strategic and International Studies, \textit{China Power: How Will the Belt and Road Initiative Advance China’s Interests?}, available at https://chinapower.csis.org/china-belt-and-road-initiative/; Andrew Chatzky \& James McBride, \textit{China’s Massive Belt and Road Initiative, Council on Foreign Relations} (May 21, 2019), available at https://www.cfr.org/backgrounder/china-massive-belt-and-road-initiative.} While there are reasons for skepticism about many claimed needs for...
protection of domestic producers, there also is special reason for wariness about the arguments certain to be made on the other side of this debate.

V. Conclusion

Given the paramount importance of national security, it is critical to examine complaints about the threats posed by China-sourced products in the ICT sector. The combination of factors—political, economic, military, and practical—that make such products especially likely to pose security threats provides strong reasons to consider U.S. actions that could counter such threats before there is significant damage to U.S. national security.

In particular, the Department of Commerce should view an investigation under Section 232 of the Trade Expansion Act as an appropriate vehicle for gathering the necessary information on the scope and shape of security threats posed by particular firms, products, or product classes and for formulating responses to those threats. Although Chinese officials would oppose an investigation and responses that it might generate, that opposition might say more about their interest in continued maintenance of conditions conducive to espionage (or at least to facilitating it when that would most serve China’s perceived national interests) than it does about the factual predicates for U.S. action. This paper does not purport to give a final answer to the question whether particular actions ultimately are the right responses, but it does support serious inquiry into threats to U.S. security from a broader set of firms and products than has been the focus of public scrutiny.
An Unconstitutional Attempt to Address Affordable Housing

By Jeffrey M. Harris

Members of Congress recently introduced the Affordable Housing Credit Improvement Act of 2019 (S.1703) to "expand and strengthen the Affordable Housing Tax Credit (also known as the Low-Income Housing Tax Credit) to produce more units of affordable housing and better serve a number of at-risk and underserved communities." Although these are important goals, the Act seeks to pursue them in a manner that disrupts decades of settled expectations, retroactively changes the terms of already-executed affordable housing partnerships, strips investors of valuable property and contract rights, and disregards foundational principles of tax law. The relevant provisions of the Act (in particular, Sections 303(b)(3) and (c)(2)) are incompatible with the Due Process and Takings Clauses of the U.S. Constitution, as well as the longstanding "economic substance" doctrine of tax law. Those provisions would likely be found unconstitutional if enacted.

I. Background

A. Overview of the LIHTC Program

In the United States, there is a significant shortage of affordable housing available to extremely low-income ("ELI") households, whose income is at or below 30% of the median income for the area. According to some estimates, there is a shortage of 7.4 million affordable rental homes for this population, which means there are only 35 units available for every 100 ELI households. Moreover, an estimated 12 million households are forced to pay over 50% of their annual income for housing. Housing instability has been shown to adversely affect employment and academic achievement, as well as physical and mental health outcomes.

The federal government has adopted multiple programs to address these important issues. Chief among them is the Low-Income Housing Tax Credit ("LIHTC"), codified in Section 42 of the 1986 Tax Reform Act. The LIHTC provides tax credits to incentivize private acquisition, development, and rehabilitation of affordable rental housing for low-income households. This program costs the federal government approximately $9 billion per year in foregone tax revenue, which makes it by far the largest federal program to address affordable housing. The LIHTC has helped to create or maintain over 2.4 million homes in

Abstract:

This paper analyzes recent legislative proposals that are purportedly designed to strengthen and improve the federal affordable housing program. The author argues that these proposals would violate multiple provisions of the U.S. Constitution and would upend bedrock principles of tax law.
The specific amount of the credit varies depending on the states and is awarded to various projects through available tax credits to the states based on their population, and affordable housing projects would not be economically feasible.

To implement the LIHTC program, Congress allocates the available tax credits to the states based on their population, and the states then award the credits to various projects through a competitive process that is overseen by each state's housing finance agency. The specific amount of the credit varies depending on whether the project is financed through tax-free bonds; projects that use tax-free bonds may receive a credit of 4% of the project's qualified basis (i.e., cost of construction), and those that do not use tax-free bonds may receive a 9% credit.

To qualify for a credit, a project sponsor must agree to set aside at least 40% of the units for renters earning no more than 60% of the area's median income or 20% of the units for renters earning 50% or less of the area's median income. These units are also subject to rent restrictions under which the maximum permissible gross rent, including an allowance for utilities, must be less than 30% of the renter's imputed income based on the area's median income. The affordability restrictions must remain in place for a minimum of 15 years.

LIHTC-financed projects are typically structured as limited partnerships between the sponsor/developer of the project—which is often a nonprofit organization that focuses on low-income housing—and an investor or group of investors. The developer serves as the general partner and exercises management authority over the project, but typically retains only a nominal equity stake (1% or less). The investors, by contrast, are limited partners who make a direct capital investment in the project in exchange for 99% or more of the project's equity. This structure allows the investors to claim the vast majority of the LIHTC tax credits, since they have provided nearly all of the equity; the tax credits are also useless to a nonprofit entity that pays no federal income tax. Although the investors typically have limited authority over the management or operation of the project, most partnership agreements require the consent of the investors before any capital event such as a sale or refinancing.

The LIHTC tax credits may be claimed over a 10-year period, beginning when the project is placed into service. But the affordability limitations and rent caps must remain in place for 15 years. After that 15-year compliance period has ended, the IRS will no longer attempt to recapture the tax credits for non-compliance.

Although the LIHTC tax credits are the primary reason why investors participate in affordable housing projects, they are by no means the only reason. In addition to the tax credits, a limited partner investor may also receive tax losses from depreciation deductions as well as residual value upon sale if the project appreciates. Indeed, it is critical for the limited partner investor to have upside potential in any appreciation as well as downside risk if the project fails; if the investor does not have upside potential and downside risk, it may be treated as a mere lender rather than an owner of the project and would thus be ineligible to claim any tax credits.

A sale or change in the ownership structure of a LIHTC project can happen at any time, but it is most likely to happen after year 15. All of the tax credits have been claimed by year 10, and those credits cannot be recaptured by the IRS after year 15. The limited partner investors may be able to obtain additional benefits from ongoing ownership after year 15 (such as depreciation deductions and other tax losses), but many seek to exit the project at that time. In many—but by no means all—projects, a limited partner investor exits the partnership by conveying its ownership interests to the general partner nonprofit. A 2012 survey of syndicators and investors found that between 60% and 85% of properties are ultimately transferred to the nonprofit.

B. Options and Rights of First Refusal in LIHTC Partnerships

The partnership agreements between general partners and limited partners typically include multiple avenues through which ownership of a project can be conveyed from the limited partners to the general partner after year 15. One common term is a purchase option. An option “gives the optionee the right to purchase the property at his election within an agreed period at a named price.” The purchase option in a LIHTC agreement will typically allow the nonprofit general partner to unilaterally purchase the property at market price within a set period of time after the end of the 15-year compliance period.

In addition to a purchase option, many LIHTC partnerships also grant the nonprofit general partner a separate right of first refusal (ROFR). Section 42(f)(7)(A) provides that:

No Federal income tax benefit shall fail to be allowable to the taxpayer with respect to any qualified low-income building merely by reason of a right of 1st refusal held by a qualified nonprofit organization to purchase the property after the close of the compliance period for a
price which is not less than the minimum purchase price determined under subparagraph (B).\textsuperscript{17} The “minimum purchase price,” in turn, is defined as the sum of “the principal amount of outstanding indebtedness secured by the building” plus “all Federal, State, and local taxes attributable to such sale.”\textsuperscript{18} In short, Section 42 allows a LIHTC project to maintain its eligibility for tax credits notwithstanding the inclusion of a ROFR for the nonprofit at a price equal to the value of outstanding debt on the property plus exit taxes. The debt-plus-taxes price under Section 42(i)(7) is typically far below the market value of the property.

Congress did not define the term “right of 1st refusal” in Section 42(i)(7). But, as the Supreme Court has repeatedly held, “[i]t is a settled principle of interpretation that, absent other indication, ‘Congress intends to incorporate the well-settled meaning of the common-law terms it uses.’”\textsuperscript{19} That is:

where Congress borrows terms of art in which are accumulated the legal tradition and meaning of centuries of practice, it presumably knows and adopts the cluster of ideas that were attached to each borrowed word in the body of learning from which it was taken and the meaning its use will convey to the judicial mind unless otherwise instructed.\textsuperscript{20}

In such cases, the “absence of contrary direction” from Congress “may be taken as satisfaction with widely accepted definitions, not as a departure from them.”\textsuperscript{21}

“Right of first refusal” is a legal term of art that has a well-established meaning at common law.\textsuperscript{22} Section 42(i)(7) must accordingly be interpreted in light of this settled understanding about the meaning of a ROFR.

A ROFR is fundamentally a defensive or preemptive mechanism that “limits the right of the owner to dispose freely of its property by compelling the owner to offer it first to the party who has the first right to buy.”\textsuperscript{23} That is, “a ‘right of first refusal’ means ‘the [r]ight to meet terms of [a] proposed contract before it is executed; e.g. right to have [the] first opportunity to purchase real estate when such becomes available, or [the] right to meet any other offer.’”\textsuperscript{24} There are several integral components of a ROFR.

1. Bona Fide Third-Party Offer

A ROFR is triggered only by a bona fide offer to purchase the property that is made by a third party. An “offer” means a proposal that “results in a binding contract upon acceptance by the other party acceding to its terms.”\textsuperscript{25} An “indefinite” proposal or mere “invitation to enter into negotiations” is insufficient.\textsuperscript{26}

Moreover, an offer must be bona fide and “made in good faith.”\textsuperscript{27} The purpose of this requirement is to provide “protection . . . against a sham offer, made not in good faith, precipitating exercise of the preemptive right.”\textsuperscript{28} Absent the requirement of a bona fide offer, a party could seek to self-trigger its ROFR by soliciting a sham offer from a friend who had no intention of actually buying the property. In determining whether an offer is bona fide, courts consider factors such as the relationship of the parties, whether the offer approximates fair market value, and whether there is any fraud or misrepresentation.\textsuperscript{29}

2. Acceptance

It is equally well established that the property must actually be for sale and that the owner must indicate a willingness to accept the third-party offer in order to trigger preemptive rights under a ROFR. That is, “the holder of a right of first refusal holds only a general contract right to acquire a later interest in real estate should the property owner decide to sell.”\textsuperscript{30} In sum, if the property owner “only received an offer on its property and did not display any desire or willingness to sell,” then “as a matter of law, the right of first refusal is not operative.”\textsuperscript{31}

3. No Power to Compel a Sale

Relatively, courts have made clear that a ROFR can never be invoked to force an unwilling property owner to sell. Instead, a ROFR is a purely defensive mechanism that is triggered only.

\textsuperscript{17} 26 U.S.C. § 42(i)(7)(A).
\textsuperscript{18} Id. § 42(i)(7)(B).
\textsuperscript{19} Sekhar v. United States, 570 U.S. 729, 732 (2013) (citing Neder v. United States, 527 U.S. 1, 23 (1999)).
\textsuperscript{20} Morissette v. United States, 342 U.S. 246, 263 (1952).
\textsuperscript{21} Id.; see also Autoria Federal Savings & Loan Ass’n v. Solimino, 501 U.S. 104, 108 (1991) (“[W]here a common-law principle is well established, . . . the courts may take it as a given that Congress has legislated with an indication, ‘Congress intends to incorporate the well-settled meaning at common law.’”).
\textsuperscript{22} See, e.g., Keeper’s, Inc. v. ATGCNG Realestate, LLC, 80 A.3d 88, 91-92 (Conn. App. Ct. 2013) (“[W]hat constitutes a ‘right of first refusal’ has been well-defined, and this term “has been defined and distinguished in many treaties and reported decisions.”).
\textsuperscript{24} Tachdjian, 568 S.E.2d at 66 (quoting Black’s Law Dictionary (6th ed. 1990)).
\textsuperscript{26} Id.
\textsuperscript{27} Id.; see also Jones v. Riley, 471 S.W.2d 650, 659 (Tex. App. Ct. 1971) (“A ‘bona fide offer’ . . . had to not only be made in good faith, but it had to also be of such a nature and in such form that it could be, by an acceptance thereof by the offeree, caused to ripen into a valid and binding contract that could be enforced by any party to it.”).
\textsuperscript{28} Steuart v. McChesney, 444 A.2d 659, 663 (Pa. 1982).
\textsuperscript{29} See DCM Inv. Corp. v. Pinecrest Inv. Co., 34 P3d 785, 789 (Utah 2001).
\textsuperscript{30} Jones v. Stahr, 746 N.W.2d 394, 399 (Neb. Ct. App. 2008) (emphasis added); see also Riley v. Camppeau Homes, Inc., 808 S.W.2d 184, 188 (Tex. App. 1991) (“A right of first refusal ripens into an option when the owner elects to sell.”); Advanced Recycling, 787 N.W.2d at 783 (“A right of first refusal ripens into an option contract when the owner receives the third-party offer and manifests an intention to sell on those terms.”); Mandell v. Mandell, 214 S.W.3d 682, 688 (Tex. App. 2007) (“A preferential right of purchase is a right granted to a party giving him or her the first opportunity to purchase property if the owner decides to sell it.”).
when a property owner decides to sell and receives a bona fide third-party offer. As a Texas appellate court explained, “[u]nlike an option contract, a right of first refusal does not give the [holder] the power to compel an unwilling owner to sell.” Furthermore:

An owner does not have to sell and, until the owner decides to sell, there is nothing to exercise, . . . However, once an owner decides to sell, there is an obligation to offer the holder of the right of first refusal the opportunity to buy the burdened property on the terms offered by a bona fide purchaser.

4. Option v. ROFR

The core attributes of a ROFR discussed above make clear that a ROFR is fundamentally different from an option. An option “gives the optionee the right to purchase the property at his election within an agreed period at a named price.” A ROFR, by contrast, is a “conditional” right that “ripen[s] into an option.” However, once an owner decides to sell, there is no obligation to offer the holder of the right of first refusal the opportunity to buy the burdened property on the terms offered by a bona fide purchaser.

5. Sale Price

There is only one way in which Congress indicated an intent to depart from the longstanding common-law concept of a ROFR. At common law, the ROFR price would be equal to the price at which the third party offered to buy the property. Section 42(i)(7), however, authorizes a ROFR at a below-market price consisting of the value of outstanding debt plus exit taxes. Thus, in this single aspect of the ROFR, Congress has expressed its intent to modify the common-law definition. In all other aspects, however, Congress has left unchanged “the well-settled meaning of the common-law term "right of first refusal".”

If ownership is not transferred to a nonprofit through exercise of a purchase option or ROFR (or some other agreement between the investors and the nonprofit), the property can then be sold pursuant to Section 42(h)(6). That provision requires the limited partner to give the state tax credit agency one year to find a qualified buyer to purchase the property at the price of debt plus taxes. If no buyer is found, the project can then be sold freely at market price. A project may also change hands through foreclosure, abandonment, or charitable contributions, but these are much less common exit strategies.

C. The Legislative History of Section 42(i)(7)

Section 42(i)(7) was added to the LIHTC program through the Low-Income Housing Tax Credit Act of 1989 (“1989 Act”). The legislative history of that provision eliminates any doubt that Congress expected the “right of 1st refusal” held by a nonprofit general partner to be an actual right of first refusal as that term is understood at common law, not just an option. Indeed, Congress made clear that a ROFR, rather than an option, was needed to avoid disrupting longstanding principles of tax law.

In May 1988, Senators George Mitchell and John Danforth convened a task force to review the operations of the LIHTC program and propose improvements. Their report, issued in January 1989, expressed concerns that affordable housing projects would be sold to for-profit entities after the expiration of the compliance period and would no longer be available for low-income tenants. The report thus urged Congress to explore new ways to ensure the ownership and management of affordable housing projects by nonprofit groups, including by granting nonprofits an option to purchase the properties at a below-market price after the end of the 15-year compliance period.

Consistent with the recommendations of the Mitchell Report, an early version of the 1989 Act proposed the use of
below-market purchase options.44 But Congress expressly rejected that proposal “because of the tax policy concern that use of such options removed any reasonable expectation that investors would derive a profit independent of tax benefits.”45 Congress was concerned that the “grant of a below-market option . . . was a substantial enough relinquishment of one of the benefits of ownership” that it would be questionable whether the investors retained a sufficient ownership interest to be eligible to receive the tax credits.46 This concern arose because of the “economic substance” doctrine of tax law, under which every business transaction must have a “substantial purpose (apart from Federal income tax effects),” and entering into the transaction must “change[ ] in a meaningful way (apart from Federal income tax effects)” the taxpayer’s economic position.”47

Instead of allowing a below-market purchase option, Congress enacted a “compromise” proposal: “a special rule that permits owners to receive the credit and other tax benefits even though the tenants hold a right of first refusal for the purchase of their units (at the end of the fifteen-year compliance period) for a specified minimum purchase price.”48 Congress determined that the use of a ROFR—which, as explained above, confers far more limited rights than an option—would still leave the limited partner investors with a sufficient economic interest to satisfy the economic substance rule.

The legislative history further shows that Congress expected the Section 42(i)(7) “right of 1st refusal” to include the standard common-law accoutrements of a ROFR. As Tracy Kane, the tax consultant to the ranking senator on the Finance Committee, explained:

The formula [for the] right of first refusal is a rather unusual legislative creation. Normally a right of first refusal is “a right to buy before or ahead of another; thus . . . the contract gives to the prospective purchaser the right to buy upon specified terms, but, and this is the important point, only if the seller decides to sell.” Therefore, unlike an option, the right of first refusal does not give the holder the power to compel an unwilling owner to sell. The compromise was most likely structured in this manner because the right of first refusal leaves more power in the hands of the owner whereas a purchase option would have given more discretion to the prospective buyer.49

These sources make clear that the ROFR contemplated by Section 42(i)(7) would not confer on the nonprofit a unilateral right to compel a sale, but would instead be triggered only after the owners had “decided to sell” and had received a bona fide third-party offer.

D. Recent Litigation and Legislative Proposals Regarding the Right of First Refusal

In the early 2000s, the first wave of affordable housing projects to be financed through the LIHTC program began reaching year 15. When a project was in a distressed area or was worth the amount of existing debt or less, the limited partner investors would generally be willing to exit the project and sell to the developer at the Section 42(i)(7) price: outstanding debt plus exit taxes.

But many projects in desirable or high-cost areas had appreciated in value and were worth considerably more than outstanding debt plus taxes. As noted, the limited partner investors were entitled to their share of such appreciation; if they had not been able to obtain this “upside potential,” then it would have been questionable as to whether they were actually owners of the property for tax purposes under the economic substance doctrine.50

For projects that had appreciated in value, some limited partners insisted on adherence to contractual terms about whether and under what conditions a nonprofit was entitled to exercise a ROFR and buy the property at the Section 42(i)(7) price. For example, the investors may have refused to allow a nonprofit to invoke its rights under a ROFR absent a bona fide, third-party purchase offer that was deemed acceptable to the partnership. These disputes have led to both litigation and proposed legislation.

1. The Memorial Drive Litigation

The Memorial Drive case involved a LIHTC partnership for a large affordable housing project in Cambridge, Massachusetts.51 As in many other agreements, the general partner possessed both an option and a ROFR. The option allowed the general partner to purchase the development at market price at any time after the 15-year compliance period had run.52 The ROFR, by contrast, would be triggered only if the general partner produced an “offer to purchase the property” from a third party, along with the price being offered and “all other terms of the proposed disposition.”53 Once the ROFR was triggered, the nonprofit would have the right to purchase the property at the lesser of the Section 42(i)(7) price, the market price, or the amount of the third-party offer.54

After the 15-year period had run, the general partner in the Memorial Drive project sought to buy out the limited partners’ interests at the Section 42(i)(7) price. But the limited partners refused that deal, arguing that the general partner could not exercise its right under the ROFR to buy at that price unless a

44 See S.B. 980, 101st Cong. (1st Sess. 1989) (providing that determination of ownership for tax purposes of a LIHTC project “shall be made without regard to any option by a qualified nonprofit organization . . . to acquire such building at less than fair market value after the close of the compliance period”).


46 Id. at 893.


48 Id., supra note 45, at 896 (emphasis added).

49 Id. at 896-97 (emphasis added); see also 136 Senate Congressional Record for Oct. 18, 1990 at S30528 (noting that ROFR rights and below-market price are available only “should the owner decide to sell”); H.R. Rep. No. 101-247 at 2665 (1989) (same).

50 See, e.g., Historic Boardwalk Hall, 694 F.3d at 454-55.


52 Id. at *4-5.

53 Id. at *2-4.

54 Id.
third-party offer had been made that was acceptable to the entire partnership. The general partner then reached out to another nonprofit and asked it to make an offer as a “favor” solely to trigger the ROFR. The investors argued that this sham offer could not trigger the ROFR because it was not a bona fide offer and the partnership had not consented to a sale.

The Massachusetts courts ruled in favor of the general partner. The Superior Court conceded that the investors’ position “has some superficial appeal” based on the language of the contract. But the court ultimately ruled in favor of the general partner based largely on what it deemed the “purpose” of Section 42. For example, the court noted that if a nonprofit needed to pay more than the Section 42(i)(7) price, this would “limit the cash flow that is available for operating the property and meeting its capital needs.” A transfer at the Section 42 price thus “contributes to the overall goal of promoting the continuing availability of affordable housing.”

The court further concluded that allowing the general partner to solicit a sham offer to trigger the ROFR would not “deprive the defendants of the benefit of their bargain” since they would still have been able to claim the available tax credits. The court asserted that “maximizing tax benefits for the Limited Partners was a key component of the arrangement,” and that there was “no language to support the claim that the Limited Partners expected to receive the residual value of the property on a sale.”

In other words, the investors had made enough money through the tax credits, so the court would not enforce the clear contractual terms. But the court ultimately ruled in favor of the general partner based largely on what it deemed the “purpose” of Section 304 of the bill proposed a “modification of rights related to building purchase” by a nonprofit. Section 304 would have expressly converted the right of first refusal authorized by Section 42(i)(7) into an option. The legislation expressly recognized that this proposal was a “modification of rights” that would change existing law.

Critically, however, this significant change to the nature of the ROFR would apply only prospectively to new projects. Section 303(c) provided that “[t]he amendments made by this section shall apply to agreements entered into or amended after the date of enactment of this Act.” The 2017 Cantwell-Hatch bill thus recognized that it would severely disrupt reliance interests and investment-backed expectations if the new modifications applied retroactively to existing projects that had already been negotiated and financed in reliance on the current state of the law. The Cantwell-Hatch bill was introduced in March 2017, but no further action was taken on it during the 115th Congress.

3. The SHAG Litigation

Meanwhile, another case about the scope of investors’ ROFR rights was progressing through the federal district court in Washington State. In _Senior Housing Assistance Group v. AMTAX Holdings 260, LLC_, a nonprofit developer (SHAG) sought to self-trigger its right of first refusal just days before it expired by soliciting a so-called offer from a friend who had no intention of actually buying the property.

The court squarely rejected that maneuver. Because the term “right of 1st refusal” in Section 42(i)(7) was undefined, the court looked to long-established common-law principles interpreting that concept. As the court explained, a right of first refusal is a legal term of art, and is triggered only if the owner receives a “bona fide offer from a third party, acceptable to the property owner.” To be “bona fide,” the offer must be “made in good faith; without fraud or deceit,” and must be sincere or genuine. The offer must also be enforceable, and not merely “an expression of interest or invitation to negotiate.”

Applying those longstanding principles, the court held that the offers at issue were insufficient to trigger SHAG’s ROFR. The offers in question were “not made in good faith and [...] not sincere or genuine,” but were instead “sham offer[s]” made by a friend of SHAG’s owner “solely as a business favor that could pay dividends in future business dealings.” Moreover, even if the offers had been bona fide, they could not trigger the ROFR because the property owner “never formed or expressed a willingness to accept” them.

Finally, the court also concluded that SHAG was not entitled to equitable relief (such as an order of specific performance on the contract terms) because it acted with unclean hands. As the court explained, SHAG “engaged in inequitable, bad faith, and unjust conduct when it secretly colluded with [a friend] to procure sham offers from straw buyers” for the projects in question.

55 Id. at *5.
56 Id. at *6.
57 Id. at *7.
58 Id.
59 Id. at *9.
60 Id.
62 See S.548, § 303(a) (amending Section 42(i)(7) by “striking ‘a right of 1st refusal’ and inserting ‘an option’”).
64 Id. at *9 (quoting Matson v. Emory, 676 P.2d 1029 (Wa. 1984)).
65 Id. at *10.
66 Id.
67 Id.
68 Id.
69 Id. at *12.
thus entered judgment for the limited partners, holding that SHAG’s ROFRs “were neither triggered nor validly exercised.”70

4. The 2019 Cantwell Bill

The SHAG court issued its decision on March 29, 2019. Two months later, on June 4, 2019, Senator Cantwell introduced a revised version of the Affordable Housing Improvement Act (S.1703). Like the 2017 version of the bill, S.1703 would convert the “right of 1st refusal” safe harbor in Section 42(i)(7) into an option, and it would apply this “modification” prospectively to new projects only.

But the 2019 bill also includes several provisions not present in the bipartisan 2017 bill that appear designed to override the holding of the SHAG court. In a section labeled “clarification with respect to right of first refusal and purchase options,” the 2019 bill would add language to Section 42(i)(7) stating:

For purposes of determining whether an option, including a right of first refusal, to purchase property is described in the preceding sentence—

(i) such option or right of first refusal may be exercised with or without the approval of the taxpayer, and

(ii) a right of first refusal may be exercised in response to any offer to purchase the property, including an offer by a related party.71

These provisions of S.1703 would effectively gut the key requirements of a right of first refusal by stripping any consent rights for limited partners and by allowing the ROFR to be triggered by a sham offer from a person who has no bona fide intent to purchase the property. The legislation would abrogate the holding of the SHAG court, which refused to allow a general partner to self-trigger its ROFR by soliciting a sham offer from a friendly party.

Furthermore, Section 303(c)(2) provides that, “[t]he amendments made by subsection (b) shall apply to agreements among the owners of the projects . . . entered into before, on, or after the date of enactment of this Act.” These so-called clarifications would thus apply retroactively to projects that were already negotiated, executed, and financed in reliance on the existing state of the law regarding rights of first refusal.

II. Analysis

A. The Act’s Retroactive Application to Already-Executed Partnership Agreements Violates the Due Process Clause

1. The Due Process Clause Limits Congress’ Ability to Pass Retroactive Laws

Our Constitution strongly disfavors retroactive lawyaking.72 The principle that laws should not apply retroactively “has long been a solid foundation of American law” and “has timeless and universal human appeal.”73 Indeed, “our law has harbored a singular distrust of retroactive statutes” for “centuries.”74 As Justice Joseph Story recognized more than 150 years ago, “retroactive laws . . . are generally unjust; and, as has been forcibly said, neither accord with sound legislation nor with the fundamental principles of the social compact.”75 In a “free, dynamic society, creativity in both commercial and artistic endeavors is fostered by a rule of law that gives people confidence about the legal consequences of their actions.”76

Retroactive legislation “presents problems of unfairness that are more serious than those posed by prospective legislation, because it can deprive citizens of legitimate expectations and upset settled transactions.”77 For example, “if retroactive laws change the legal consequences of transactions long closed, the change can destroy the reasonable certainty and security which are the very objects of property ownership.”78 Thus, “whereas prospective economic legislation carries with it the presumption of constitutionality, it does not follow . . . that what Congress can legislate prospectively it can legislate retrospectively.”79

Although retroactive legislation implicates a number of constitutional provisions—including the Takings Clause, Ex Post Facto Clause, Bill of Attainder Clause, and others—the primary protection against “retroactive laws of great severity”80 lies in the Due Process Clause. The Due Process Clause gives effect to the law’s general distrust of retroactive laws by “protect[ing] the interests in fair notice and repose that may be compromised by retroactive legislation.”81 “The retrospective aspects of [economic] legislation, as well as the prospective aspects, must meet the test of due process, and the justifications for the latter may not suffice for the former.”82 The Supreme Court has accordingly “treat[ed] due process challenges based on the retroactive character” of legislation “as serious and meritorious, thus confirming the vitality of our legal tradition’s disfavor of retroactive economic legislation.”83 “Both stability of investment and confidence in the constitutional system . . . are secured by due process restrictions against severe retroactive legislation.”84

70 Id. at *13.
71 S.1703, § 303(b)(3).
74 Eastern Enterprises, 524 U.S. at 547 (Kennedy, J., concurring in the judgment and dissenting in part).
75 Id. (quoting 2 Joseph Story, Commentaries on the Constitution § 1398 (2d ed. 1851)).
76 Landgraf v. USI Film Prods., 511 U.S. 244, 266 (1994).
78 Eastern Enterprises, 524 U.S. at 548 (Kennedy, J.).
79 Id. at 547-48 (quoting Usery v. Turner Elkhorn Mining Co., 428 U.S. 1, 15 (1976)).
80 Id.
81 Landgraf, 511 U.S. at 266.
82 Eastern Enterprises, 524 U.S. at 548 (Kennedy, J.).
83 Id.
84 Id. at 549.
In determining whether legislation operates retroactively, the Supreme Court has drawn on an “influential definition” offered by Justice Story in 1814. Under that definition, a statute is retroactive if it “takes away or impairs vested rights acquired under existing laws, or creates a new obligation, imposes a new duty, or attaches a new disability, in respect to transactions or considerations already past.” In short, “the court must ask whether the new provision attaches new legal consequences to events completed before its enactment.” That inquiry considers “the nature and extent of the change in the law and the degree of connection between the operation of the new rule and a relevant past event,” and “familiar considerations of fair notice, reasonable reliance, and settled expectations offer sound guidance.” Courts are also more likely to find impermissible retroactivity when a statute affects “substantive rights, liabilities, or duties,” as opposed to a procedural or jurisdictional rule.

2. S.1703 Would Have Significant Retroactive Effects That Violate Due Process

Although S.1703 labels the changes in Section 303(b) as mere “clarifications,” they unquestionably affect “substantive rights, liabilities, or duties.” “Right of first refusal” is a term of art with a well-established meaning at common law. A party's purchase rights are triggered only by a bona fide third-party offer, and only when the owner has decided to sell the property and has indicated a willingness to accept the offer. Yet Section 303(b)(3) would eliminate these aspects of the ROFR. Section 303(b)(3) provides that a ROFR may be exercised “with or without the approval of the [investor].” Section 303(b)(3) would also have a “severe” impact on the “stability of investment,” because the ROFR that this legislation seeks to modify is an integral component of LIHTC partnership agreements. When a ROFR is triggered, it allows a nonprofit to buy out its limited partner investors at a price far below the fair market value of the building, thereby allowing the nonprofit to effectively capture all of a project’s appreciation in value.

The precise scope of a ROFR, and the conditions under which it can be triggered, are thus of central importance to a LIHTC partnership agreement and have powerful and far-reaching implications for both the investors and the nonprofits. Congress recognized as much when it enacted Section 42(i)(7) in 1989. As explained above, the legislative history of this provision shows that Congress was well aware when it enacted Section 42(i)(7) that “the right of first refusal does not give the holder the power to compel an unwilling owner to sell,” and that the enacted version “leaves more power in the hands of the owner” than a purchase option would. Yet S.1703 now seeks to retroactively undo this careful legislative compromise. This attempt to readjust core provisions

---

85 Landgraf, 511 U.S. at 268.
86 Id. at 269 (emphasis added).
87 Id. at 269-70.
88 Id.; see also id. at 280 (asking whether legislation “would impair rights a party possessed when he acted, increase a party's liability for past conduct, or impose new duties with respect to transactions already completed”).
89 Id. at 274-78.
90 Id. at 278 (emphasis added).
91 See SHAG, 2019 WL 1417299 at *9-10.
92 Steuart, 444 A.2d at 663.
93 Landgraf, 511 U.S. at 271.
94 Id. at 269 (emphasis added).
95 See, e.g., Eastern Enterprises, 524 U.S. at 549 (Kennedy, J.) (legislation that “creat[ed] liability for events which occurred 35 years ago . . . has a retroactive effect of unprecedented scope”).
96 Id.
97 See Kaye, supra note 45, at 895-96.
of transactions negotiated decades ago is "far outside the bounds of retroactivity permissible under our law." 98

Section 303(b)(3) also appears designed to abrogate the holding of the SHAG case—decided just a few months before S.1703 was introduced—which properly enforced the conditions of a ROFR and held that a nonprofit could not self-trigger its ROFR by soliciting a sham offer as a favor from a friend. Section 303(b)(3) would gut the holding of SHAG by eliminating any need for the limited partner to accept an offer, and by allowing any offer (even a sham offer solicited by the nonprofit) to trigger the ROFR. The fact that Section 303(b)(3) appears designed to override the holding of a decision by a federal district court is a powerful indication that the legislation is not merely clarifying the law but is instead attempting to retroactively adjust substantive rights.

Proponents of Section 303(b)(3) may argue that this legislation is not retroactive because it merely affects the tax treatment of ROFRs rather than directly abrogating contract or property rights. They will likely argue that Section 42(i)(7) does not create ROFRs or require ROFRs but instead merely creates a safe harbor under which the use of a ROFR as specified in that section will not jeopardize a project’s eligibility for LIHTC credits. But that argument ignores how LIHTC projects work in practice. Section 42 is referenced in many LIHTC partnership agreements, so courts will often look to Section 42 in interpreting the scope of ROFR rights under a partnership agreement. For example, the limited partnership agreements at issue in the SHAG case repeatedly referenced Section 42(i)(7). The partnership agreement at issue in Memorial Drive similarly contained a clause stating that the parties “wish to enter into a right of first refusal agreement with respect to the Property “in accordance with Section 42(i)(7) of the Internal Revenue Code.” Thus, any modification of the meaning of “right of 1st refusal” under Section 42(i)(7) will likely have an equivalent impact on the courts’ interpretation of ROFR rights under LIHTC partnership agreements. And the impact of such a holding will be to severely disrupt existing property and contract rights arising out of transactions that were negotiated decades ago.

Proponents of Section 303(b)(3) may also argue that this legislation is a proper “clarification” of the law because it ensures that ROFR rights have some value to the nonprofits. According to this line of reasoning, which the Massachusetts state courts endorsed in the Memorial Drive litigation, a proper application of the common-law requirements of bona fide offer and acceptance would mean that the nonprofits’ ROFRs were “almost never triggered” because third parties would be unlikely to seek to purchase a property on which another party held a below-market ROFR. 99 But, to the contrary, the ROFR remains highly valuable to a nonprofit even if it is enforced consistent with its common-law meaning. Most importantly, the ROFR ensures that the nonprofit will be able to remain in the partnership by limiting investors’ ability to convey the property to an outside party. That is, regardless of how often ROFR rights are actually triggered, they remain highly valuable to the nonprofit as a defensive or preemptive mechanism to ensure that the nonprofit remains involved in the operation and management of the project, and that the nonprofit cannot be cut out of the project without first being given a chance to buy the property at a favorable price. There is accordingly nothing anomalous about insisting that a ROFR in a LIHTC partnership agreement be interpreted consistent with all applicable common-law requirements.

3. At a Minimum, S.1703 Does Not Speak with Sufficient Clarity to Have Retroactive Effect

For all the reasons discussed above, Section 303(b)(3) is no mere clarification of the law but instead a substantive modification of existing contract and property rights that would violate the Due Process Clause’s prohibition on legislation with severe retroactive effects. At a minimum, however, this legislation is insufficiently clear about its intent to apply retroactively.

The Supreme Court has long recognized a “presumption against retroactive legislation” that stems from “[e]lementary considerations of fairness.” 100 This presumption “is deeply rooted in our jurisprudence, and embodies a legal doctrine centuries older than our Republic.” 101 Where the presumption applies, “congressional enactments” are not “construed to have retroactive effect unless their language requires this result.” 102 This is the case “[e]ven where some substantial justification for retroactive rulemaking is presented.” 103

Here, far from speaking clearly about the intent and effect of Section 303(b)(3), the Act describes the change it seeks to make as a mere “clarification” of the law rather than what it actually is: a major change to the contractual rights of thousands of investors in affordable housing projects. This legislation would not only destroy valuable contract and property rights but also abrogate the holding of a major decision from a federal court. If the statute is to be interpreted to effectuate these major retroactive changes to the law, Landgraf makes clear that Congress must speak with greater clarity than it has done in S.1703.

B. The Act Would Strip Limited Partner Investors of Valuable Property and Contract Rights Without a Public Purpose or Just Compensation, in Violation of the Takings Clause

Even if Section 303(b)(3) could satisfy the Due Process Clause, and even if the legislation were clear and candid about its intent to make substantive changes in the law retroactively, it would separately violate the Takings Clause, which provides that “private property [shall not] be taken for public use, without just compensation.” 104

1. The Supreme Court’s Takings Jurisprudence

The Supreme Court has “been unable to develop any ‘set formula’ for determining when ‘justice and fairness’ require that

100 Landgraf, 511 U.S. at 265.
101 Id.
102 Bowen, 488 U.S. at 208 (emphasis added).
103 Id.
104 U.S. Const., amdt. V.
economic injuries caused by public action be compensated by the government."\textsuperscript{105} Instead, the Court "has examined the 'taking' question by engaging in essentially ad hoc, factual inquiries that have identified several factors . . . that have particular significance."\textsuperscript{106} "The general rule at least is that while property may be regulated to a certain extent, if regulation goes too far it will be recognized as a taking."\textsuperscript{107}

It is well established that investors' contractual rights are private property protected by the Takings Clause.\textsuperscript{108} The three primary factors courts consider in determining whether a taking has occurred are: (1) "the economic impact of the regulation on the claimant"; (2) "the extent to which the regulation has interfered with distinct investment-backed expectations"; and (3) "the character of the governmental action."\textsuperscript{109} In applying this test, the Supreme Court has been guided by the principle that the Takings Clause "was designed to bar Government from forcing some people alone to bear public burdens which, in all fairness and justice, should be borne by the public as a whole."\textsuperscript{110}

2. S.1703 Would Effect Takings of the Property of Limited Partner Investors in the LIHTC Program

All three "primary factors" in the takings analysis support the conclusion that Section 303(b)(3) of S.1703 goes "too far" and effects an unconstitutional taking.\textsuperscript{111} First, the "economic impact" of the Act on limited partner investors would be severe. The Act would strip investors of a highly valuable contractual right that forms an essential part of the complex contracts between limited partners and general partners in LIHTC developments. By modifying critical aspects of the ROFR to allow a nonprofit to self-trigger its purchase right by soliciting a sham offer from a related party, Section 303(b)(3) would effectively ensure that limited partner investors are never able to share in the upside potential of their projects.

This legislation would also nullify the blocking or consent rights that many investors demanded as a precondition to putting millions of dollars of equity into a project. Indeed, most investors never would have accepted a below-market ROFR at all but for the protection provided by their blocking or consent rights for any capital events. Under the modified version of the ROFR contemplated by Section 303(b)(3), however, a nonprofit would be able to unilaterally force a transfer to itself at a below-market price notwithstanding clear language in the partnership agreement requiring the investors' consent before any capital event such as a sale or refinancing.

Simply put, the statute would "force[] a considerable financial burden" upon the investors.\textsuperscript{112} That economic burden, moreover, is wholly divorced from any "responsibilities that [the investors] accepted."\textsuperscript{113} The limited partners in LIHTC projects—like any equity investors—had every reason to believe based on the economic substance doctrine that they were entering into transactions in which they would receive, not just tax benefits, but meaningful upside potential if the projects were successful. And they had every reason to believe that their consent rights for capital events such as a sale of the property would be enforced as written.

Second, the Act would significantly interfere with distinct, investment-backed expectations. Whether a statute impermissibly interferes with distinct investment-backed expectations is "an objective test," because "to support a claim for a regulatory taking, an investment-backed expectation must be 'reasonable.'"\textsuperscript{114} "A reasonable investment-backed expectation must be more than a unilateral expectation or an abstract need."\textsuperscript{115} The Takings Clause "protects private expectations to ensure private investment," and "reasonable, investment-backed expectations" are to be "understood in light of the whole of our legal tradition."\textsuperscript{116} The expectations inquiry—like the due process inquiry—is animated by concerns about fair notice and reliance; its purpose is to provide compensation to "property owners who . . . bought their property in reliance on a state of affairs that did not include the challenged regulatory regime."\textsuperscript{117}

Section 303(b)(3) of S.1703 is a paradigmatic example of a statute that would effect a taking by destroying the affected parties' reasonable and distinct investment-backed expectations. The limited partner investors' contractual rights are unquestionably "property" for purposes of the Takings Clause.\textsuperscript{118} And their "expectations"—that they will continue to possess their

\textsuperscript{108} See, e.g., U.S. Trust Co. of N.Y. v. New Jersey, 431 U.S. 1, 19 n.16 (1977) ("[C]ontract rights are a form of property" for purposes of the Takings Clause.); Lynch v. United States, 292 U.S. 572, 579 (1934) ("The Fifth Amendment commands that property be not taken without just compensation to "property owners who . . . bought their property in reliance on a state of affairs that did not include the challenged regulatory regime.").
\textsuperscript{109} Connolly v. Pension Benefit Guaranty Corp., 475 U.S. 211, 224 (1986) ("The fact that legislation disregards or destroys existing contractual rights" may "transform the regulation into an illegal taking."); City of El Paso v. Simmons, 379 U.S. 497, 533-34 (1965) (Black, J., dissenting) ("Contractual rights, this Court has held, are property, and the Fifth Amendment requires that property shall not be taken for public use without just compensation.").
\textsuperscript{110} Connolly, 475 U.S. at 225 (citing Penn Central, 438 U.S. at 124).
\textsuperscript{112} Eastern Enterprises, 524 U.S. at 529-30.
\textsuperscript{113} Id. at 531.
\textsuperscript{114} Cienega Gardens v. United States, 331 F.3d 1319, 1346 (Fed. Cir. 2003) (citing Ruckelshaus v. Monsanto Co., 467 U.S. 986, 1005 (1984)); see also Colony Cove Properties, LLC v. City of Carson, 888 F.3d 445, 452 (9th Cir. 2018) ("To form the basis for a taking claim, a purported distinct investment-backed expectation must be objectively reasonable.").
\textsuperscript{115} Ruckelshaus, 467 U.S. at 1005.
\textsuperscript{116} Lucas, 505 U.S. at 1304-35 (Kennedy, J., concurring).
\textsuperscript{117} Cienega Gardens, 331 F.3d at 1345-46.
\textsuperscript{118} See, e.g., U.S. Trust Co. of N.Y., 431 U.S. at 19 n.16.
contractual rights unless and until they decide to sell or relinquish them—are clearly “investment-backed,” in that they grow out of sizeable investments made in affordable housing projects. The Act would burden those “distinct investment-backed expectations” by stripping LIHTC contracts of any meaningful constraints on the exercise of ROFR rights, notwithstanding that these agreements were negotiated in direct reliance on the then-existing state of the law (including the economic substance doctrine). The end result in effectively every LIHTC project—including those negotiated decades ago—would be to allow the general partner to self-trigger its ROFR and capture every cent of appreciation, even though the limited partner investors put up 99% or more of the capital for the project.

Moreover, Section 303(b)(3) would allow the nonprofits to unilaterally trigger their ROFRs notwithstanding any contractual consent or blocking rights held by the limited partner investors. But without the protection provided by those rights—which are especially critical in the context of agreements that include below-market ROFRs—many investors would have never committed millions of dollars of capital to LIHTC projects, or would have insisted on more generous terms before investing. By nullifying these critical contractual protections that were a precondition to multi-million-dollar investment decisions, Section 303(b)(3) would “destroy [those] previously existing rights of . . . contract.” As in Pennsylvania Coal, enacting S.1703 would “so frustrate distinct investment-backed expectations as to amount to a taking.”

The Supreme Court invoked similar principles in Eastern Enterprises v. Apfel, which held that provisions of the Coal Industry Retiree Health Benefit Act violated the Takings Clause because they “substantially interfere[d]” with “reasonable investment-backed expectations.” The statute at issue imposed significant retroactive liability on Eastern Enterprises, which was forced to pay $50 to $100 million into a health benefit fund for coal miners based on conduct that occurred “some 30 to 50 years before the statute's enactment, without any regard to responsibilities Eastern accepted under any benefit plan the company itself adopted.”

The Court concluded that the statute “substantially interfere[d] with Eastern's reasonable investment-backed expectations” because “the extent of Eastern's retroactive liability” under the law was “substantial,” “particularly far reaching,” and not “justified.” In particular, the statute impermissibly “attach[e] new legal consequences” to a “relationship completed before its enactment.” Moreover, “[t]he distance into the past that the Act reaches back to impose a liability on Eastern and the magnitude of that liability raise substantial questions of fairness.” This severe retroactive liability was far out of line with any reasonable expectations Eastern might have had, because the provisions were “not calibrated either to Eastern's past actions or to any agreement—implicit or otherwise—by the company.” The Court thus held that the “Constitution [did] not permit” a scheme that so severely interfered with Eastern's reasonable and distinct investment-backed expectations.

The same reasoning that led the Court to find a taking in Eastern Enterprises applies with full force to Section 303(b)(3) of S.1703. The “extent of [investors'] retroactive liability” under Section 303(b)(3) would be “substantial” because its enactment would eliminate valuable contractual rights that grow out of

119 See Penn Central, 438 U.S. at 127.
120 Lucas, 505 U.S. at 1015; Penn Central, 438 U.S. at 127.
121 Eastern Enterprises, 524 U.S. at 532.
122 260 U.S. at 413; see Penn Central, 438 U.S. at 127 (citing Pennsylvania Coal as “the leading case” for the proposition that a statute “may so frustrate distinct investment-backed expectations as to amount to a taking”).
124 Penn Central, 438 U.S. at 127.
125 Pennsylvania Coal, 260 U.S. at 413.
126 Penn Central, 438 U.S. at 127.
127 524 U.S. at 532.
128 Id. at 531.
129 Id. at 532, 534-35.
130 Id. at 532 (quoting Landgraf, 511 U.S. at 270).
131 Id. at 534.
132 Id. at 536.
133 Id.
134 Id. at 532.
long-settled transactions, and it would strip the limited partner investors of any upside potential from their investments or any consent rights for capital events. As in Eastern Enterprises, moreover, the Act “attaches new legal consequences” to contractual agreements “completed before its enactment.”135 And, with respect to the reasonableness of investors’ expectations, Section 303(b)(3)’s elimination of the limits on the exercise of ROFR rights is “not calibrated” to “any agreement—implicit or otherwise”—that the investors accepted when they entered the project.136 To the contrary, the investors had every reason to believe that ROFRs would be applied consistent with the longstanding, common-law meaning of that term, and that the investors would: (1) retain upside potential in their projects if a ROFR was not exercised and (2) retain their highly valuable contractual rights to consent to any disposition of the property. As in Eastern Enterprises, S.1703 “imposes[] a disproportionate and severely retroactive burden upon” investors, and “the Constitution does not permit” such a law to stand.137

Finally, the character of the governmental action also raises serious Takings Clause concerns. S.1703 would “single out” investors “to bear a burden that is substantial in amount [and] . . . unrelated to any commitment the [investors] made or to any injury they caused.”138 If the government requires certain individuals to forfeit their property rights for the benefit of the public, “the governmental action implicate[s] fundamental principles of fairness underly[ing] the Takings Clause.”139 After all, the fundamental purpose of the Takings Clause is to “bar Government from forcing some people alone to bear public burdens which, in all fairness and justice, should be borne by the public as a whole.”140 If Congress wants to help nonprofit groups raise serious money to build affordable housing and provide more resources and stronger protections for at-risk groups.”141 But even if that is a valid public purpose in the abstract, the Takings Clause still requires a fit between ends and means. Even if it is acting in pursuit of legitimate policy goals, the government may not “forc[e] some people alone to bear public burdens which, in all fairness and justice, should be borne by the public as a whole.”142 A “strong public desire to improve the public condition is not enough to warrant achieving that desire by a shorter cut than the constitutional way of paying for the change.”143 Even if the proponents of S.1703 could establish that it serves a legitimate public purpose, the government would still be required to pay just compensation to investors for the destruction of their valuable property and contract rights, which would be well into the hundreds of millions of dollars.

C. The Act’s Purported Clarification Regarding the Scope of ROFRs Is Incompatible with the Longstanding Economic Substance Doctrine

Wholly apart from the constitutional defects discussed above, the “clarification” in Section 303(b)(3) would also contravene longstanding principles of tax law on which the business community has relied for decades.

The “economic substance” doctrine is a foundational principle of U.S. tax law. Under the economic substance doctrine, a transaction that has no economic substance apart from its tax implications will not be eligible for the claimed tax benefits.145 A transaction will be deemed to have economic substance only when it 1) “has a substantial purpose (apart from Federal income tax effects),” and 2) “changes in a meaningful way (apart from Federal income tax effects) the taxpayer’s economic position.”146 Thus, when two partners enter into a business arrangement, each must have some purpose for engaging in the transaction apart from any tax benefits if they are to be eligible for any tax benefits. That is, both partners must have “really and truly intended to join

135 Id.
136 Id. at 536.
137 Id.
138 Id. at 537.
139 Id.
140 Armstrong, 364 U.S. at 49; see also Palazzolo, 533 U.S. at 617–18; Lingle, 544 U.S. at 537; Ark. Game & Fish Comm’n, 568 U.S. at 31.
141 Pennsylvania Coal, 260 U.S. at 415.
142 Kelo v. City of New London, Conn., 545 U.S. 469, 477 (2005); see also Thompson v. Consolidated Gas Corp., 300 U.S. 55, 80 (1937) (“To be sure, the Court’s cases have repeatedly stated that ‘one person’s property may not be taken for the benefit of another private person without a justifying public purpose, even though compensation be paid.’”).
144 Armstrong, 364 U.S. at 49.
145 Pennsylvania Coal, 260 U.S. at 416.
Two critical indicia of whether a person is actually a partner in a business enterprise (rather than a mere lender) are whether there is “meaningful downside risk” and “meaningful upside potential.”

Those considerations help inform whether the party has “a true interest in profit and loss,” thereby making it eligible for the tax benefits to which owners of property are entitled. For example, in Historic Boardwalk Hall, the Third Circuit held that an investor was not eligible to claim historic preservation tax credits as the owner of a project when another company held an option to purchase the property in question at a below-market price. As the court explained, “although in form [the investor] had the potential to receive the fair market value of its interest . . . in reality [the investor] could never expect to share in any upside.” That was because, “[e]ven if there were an upside,” the option holder “could exercise its Consent Option, and cut [the investor] out by paying a purchase price unrelated to any fair market value.”

Section 303(b)(3) of S.1703 flouts the economic substance doctrine. By retroactively removing the core requirements for triggering a ROFR, it would leave the nonprofits with an option to unilaterally buy out their limited partner investors at a below-market price. This would mean that the nonprofit effectively owns all of the non-tax economic value of the project—including any appreciation in the property’s value—and the investors’ participation would be for the sole purpose of claiming tax credits and depreciation losses. A limited partner with no upside potential is a partner in name only, according to the economic substance doctrine, and is therefore not entitled to the tax benefits sought.

Furthermore, current tax regulations provide that “losses, deductions, or credits” arising from a LIHTC project “may be limited or disallowed under other provisions of the Code or principles of tax law,” including the “‘sham’ or ‘economic substance’ analysis.” Thus, if the “clarification” in Section 303(b)(3) were to become law, it could jeopardize investors’ ability to claim LIHTC tax credits and thereby chill much-needed investment in such projects.

Finally, S.1703’s disregard for the economic substance doctrine would also make the legislation vulnerable to another type of takings claim. In Lucas v. South Carolina Coastal Council, the Supreme Court held that a per se taking of property occurs when a regulation “denies all economically beneficial or productive uses” of the property. If Section 303(b)(3) is enacted, it would effectively strip limited partner investors of any upside potential in their LIHTC projects. The end result would be that the sole value lies in the investors’ ability to obtain various tax benefits. But the economic substance doctrine is clear that merely entering into a transaction for tax purposes does not constitute an economically beneficial or productive use of property; the taxpayer must instead have some non-tax-related purpose for engaging in the transaction. Under Lucas, the investor would suffer a taking of its property because it would be left with no “economically beneficial or productive uses” of its ownership interest apart from tax consequences.

Congress made a deliberate policy choice in 1989 to use a ROFR rather than a purchase option in Section 42(ii)(7) precisely to avoid creating an arrangement that would violate the economic substance doctrine. Yet Section 303(b)(3) of S.1703 now seeks to achieve the exact same outcome that Congress rejected in 1989 through a “clarification” that would in fact rewrite the statute retroactively. Congress had it right in 1989. Section 303(b)(3) would force limited partner investors into transactions that lack any economic substance apart from their tax implications, thereby running headlong into the economic substance doctrine and jeopardizing investors’ ability to claim LIHTC tax credits. This legislation fails to comport with longstanding, foundational principles of tax law, and should not be enacted for this reason in addition to its constitutional infirmities.

III. Conclusion

Sections 303(b)(3) and (c)(2) of S.1703 would make significant changes to the terms of complex partnership agreements that were negotiated and finalized decades ago. This would retroactively strip investors of their valuable contract and property rights in violation of the Due Process and Takings Clauses of the U.S. Constitution, and it would leave the investors in business ventures that lack any non-tax-related economic substance, threatening the tax benefits to which they would otherwise be entitled. If enacted, S.1703 would almost certainly face meritorious constitutional challenges from the entities whose property has been taken without due process or just compensation.

---


149 Historic Boardwalk Hall, 694 F.3d at 454-55.

150 Id.

151 Id. at 460.

152 Id.; see also Ronald A. Shellan, Thinking About Year Fifteen of a Low-Income Housing Tax Credit Partnership, J. Affordable Housing & Community Development at 94, 97 (Fall 2001) (noting that “IRS personnel in private discussions have indicated that they view a right of first refusal as different from an option and may well attack an option as being outside the safe-harbor provisions found in section 42(ii)(7)”).

153 26 C.F.R. §1.42-4(b).

154 505 U.S. at 1015-16.
Credentials Not Required: Why an Employee’s Significant Religious Functions Should Suffice to Trigger the Ministerial Exception

By Nathaniel M. Fouch, Erik Money, & Thomas C. Berg

Religious Liberties Practice Group

About the Authors:
Nathaniel M. Fouch and Erik Money are third-year law students at the University of St. Thomas School of Law (Minnesota). Thomas C. Berg is the James L. Oberstar Professor of Law and Public Policy at the University of St. Thomas School of Law. Portions of this article are drawn from an amicus brief drafted by the authors for the University of St. Thomas Religious Liberty Appellate Clinic on behalf of several religious and civil liberties organizations in support of the certiorari petition in Our Lady of Guadalupe School v. Morrissey-Berru (see https://www.supremecourt.gov/DocketPDF/19/19-267/117497/20190930154315350_19-267acChristianLegalSociety.pdf). Thanks to Kim Colby and Reed Smith, co-counsel on the brief, for their input and contributions. This article reflects the views of the authors alone.

Note from the Editor:
The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the authors. Whenever we publish an article that advocates for a particular position, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

Other Views:

The ministerial exception protects religious organizations from lawsuits by their ministers for employment discrimination and other alleged employment-related wrongs. The U.S. Supreme Court unanimously affirmed the exception as a requirement of the First Amendment’s Religion Clauses in Hosanna-Tabor Evangelical Lutheran Church & School v. EEOC. The Court ruled that holding a church liable for firing or refusing to hire a minister “deprive[s] the church of control over the selection of those who will personify its beliefs,” in violation of the Free Exercise Clause, and creates “government involvement in such ecclesiastical decisions,” in violation of the Establishment Clause. The exception protects religious organizations’ internal governance and allows them to raise a defense at early stages of a lawsuit, even though nondiscrimination laws are “neutral laws of general applicability” enacted for secular purposes. Hosanna-Tabor also held that the employee who had sued—a teacher in a religiously affiliated school—was a “minister,” but it based that ruling on the case’s specific facts and left open how far the definition of minister should extend.

Since Hosanna-Tabor was decided in 2012, lower courts have divided over whether the category of minister includes teachers in religious schools who have significant religious functions or responsibilities: teaching religion classes, leading prayers or liturgies, or integrating religion into ordinary subjects. Several early decisions ruled that such teachers were ministers. But three recent decisions from the Ninth Circuit and the California Court of Appeal have ruled the other way, holding that the teachers in question fell outside the definition of minister because they lacked some sort of ministerial training, title, or other credential accompanying the religious functions. Determining who is a minister based on credentials rather than function would significantly reduce protection for religious organizations in choosing who will lead, teach, and preach their faiths. The U.S. Supreme Court has granted certiorari to review the two Ninth Circuit decisions.

The Court should rule that function is sufficient. An employee who performs a significant religious function is a

---

2 Id. at 188-89. “Church” in this context encompasses religious organizations broadly.
3 Id. at 189-90 (distinguishing Employment Division v. Smith, 494 U.S. 872, 879 (1990)).
4 Id. at 190-94.
5 See infra part I (describing the divided case law).
minister whom the organization has a right to choose without interference by courts and juries. When such a function is present and the organization holds the employee out as performing the function, then the employee's minister status should not be defeated because the employee lacks ministerial credentials such as title, training, or ordination. Treating religious function as sufficient serves the purposes of the ministerial exception and avoids evils the First Amendment was meant to prevent. There is strong evidence that this best comports with the original understanding of the Religion Clauses.

I. Case Law
A. Hosanna-Tabor

*Hosanna-Tabor* was a lawsuit by Cheryl Perich, a fourth-grade teacher at a Lutheran school. She was a “called” teacher: one who had received theological training and been commissioned by the congregation sponsoring the school. After taking medical leave for narcolepsy, Perich attempted to return to work, but the school had hired another teacher. When Perich threatened to sue, the congregation withdrew her “call” for failing to follow church conciliation procedures. The EEOC sued on her behalf for disability discrimination.

When the case reached the Supreme Court, the Justices unanimously held that the ministerial exception is required by both Religion Clauses. The Court then found that Perich was a minister within the exception, but it declined to adopt any single “rigid formula for deciding when an employee [so] qualifies,” noting that this was its “first case involving the ministerial exception.” Instead, it determined that Perich was a minister, based on “all the circumstances of her employment,” four of which the Court discussed.

The first circumstance the Court noted was Perich’s job title: “Minister of Religion, Commissioned.” Second, her title reflected a significant degree of religious training followed by a formal process of commissioning. To become a called teacher, Perich had completed eight college theology courses, obtained the endorsement of her local synod, and passed an oral examination by a faculty committee at a Lutheran college. Third, Perich “held herself out” as a minister by accepting the formal call to religious service, claiming the minister’s parsonage exclusion on her income taxes, and labeling herself a minister in her communications with the synod. Fourth, her job had significant religious duties. In addition to teaching general school subjects, she taught religion exercises, attending weekly chapel services, and leading the chapel service about twice a year.

Three Justices concurred to explain that the circumstances identified by the majority were by no means all required: that is, the definition of minister should extend more broadly than reliance on all four circumstances might suggest. Justice Samuel Alito’s concurrence, joined by Justice Elena Kagan, emphasized that reliance on ministerial title, specifically, was improper because it could discriminate against minority religions that use unfamiliar titles or none at all. To preserve equality among faiths, Alito and Kagan said, the primary criterion for a minister should be significant religious functions. Justice Clarence Thomas, concurring separately, emphasized that courts should defer to the religious organization’s good faith determination of who its ministers are. This focus, he argued, would avoid both inequality among denominations and judicial second-guessing of (i.e. entanglement in) religious organizations’ ecclesiastical decision-making.

B. Lower Courts: Functions Versus Credentials

By declining for the present to adopt a definitive test for who counts as a minister, the Supreme Court left development of the matter to lower courts, which have now divided over how to determine an employee’s minister status. Is function a dominant consideration, a sufficient but unnecessary condition, or merely one factor to be considered among others?

Numerous decisions before *Hosanna-Tabor* held that function was key in determining whether an employee was a minister. In *Rayburn v. General Conference of Seventh Day Adventists*, for example, the Fourth Circuit stated that applicability of the ministerial exception should turn on the “function of the position,” not on ordination status. The Ninth Circuit and the California courts were among those that, before *Hosanna-Tabor*, took a functional approach to defining “minister.”

After *Hosanna-Tabor*, courts continued to place heavy weight on an employee’s religious function. In *Fratello v. Archdiocese of New York*, the Second Circuit held that a school principal’s lawsuit was barred notwithstanding her secular title of “lay principal.” The court said that “the most important consideration in this case is whether, and to what extent, the plaintiff performed

---

7 565 U.S. at 177.
8 Id. at 179.
9 Id. at 190.
10 Id. at 190-92.
11 Id. at 191.
12 Id.
13 Id. at 191-92.
14 Id. at 192.
15 Id. at 198 (Alito, J. concurring).
16 Id. at 196 (Thomas, J. concurring).
17 Id. at 203-04 (Alito, J. concurring).
18 772 F.2d 1164, 1168 (4th Cir. 1985).
19 Alcazar v. Corp. of Catholic Archbishop of Seattle, 627 F.3d 1288, 1291 (2010). See also Schmoll v. Chapman University, 70 Cal. App. 4th 1434, 1439 (1999) (stating that the applicability of the ministerial exception “does not depend on the title given to the employee; rather, the determinative factor is the function of the person’s position”); Henry v. Redhill Evangelical Lutheran Church, 201 Cal. App. 4th 1041, 1053-54 (2011) (“The exception encompasses all employees of a religious institution, whether ordained or not, whose primary functions serve its spiritual and pastoral mission.”).
20 863 F.3d 190, 207, 210 (2d Cir. 2017).
important religious functions for her religious organization.” In the Seventh Circuit followed suit, holding a teacher to be a minister in Grussgott v. Milwaukee Jewish Day School. Even though only two of the four circumstances noted in Hosanna-Tabor weighed in favor of defendants, Grussgott held that the plaintiff’s religious function “outweighed” the other factors. The court did say, however, that “ministers” should be defined “case-by-case,” and it declined to make function alone the “determining” factor.

The Fifth Circuit also relied on function in Cannata v. Catholic Diocese of Austin. The court found that the plaintiff was a minister because he “played an integral role in the celebration of Mass” and, “by playing the piano during services, furthered the mission of the church and helped convey its message to the congregants.” The court viewed this as an “important function during the service.”

However, recent cases from the Ninth Circuit and California have departed from the function focus. These decisions require that the employee have significant ministerial credentials—title, training, or ordination—in addition to important religious functions to qualify as a minister. The most striking of these decisions is Morrissey-Berru v. Our Lady of Guadalupe School, where the Ninth Circuit ruled that a teacher in a Catholic school was not a minister even though she had “significant religious responsibilities.”

Those responsibilities included teaching a religion class (involving Catholic doctrine) every year, “incorporating Catholic values and teachings into her curriculum, . . . le[ading] her students in daily prayer, . . . liturgy planning for a monthly Mass, and direct[ing] and produc[ing] a performance by her students during the School’s Easter celebration every year.” These functions were insufficient, the court said, because of other factors mentioned in Hosanna-Tabor: her title was secular, she lacked any “religious credential, training, or ministerial background,” and she did not hold herself out as a minister.

Morrissey-Berru relied on the Ninth Circuit’s decision in Biel v. St. James School, the first to require credentials as well as function. Biel involved a fifth-grade teacher who taught all academic subjects (including a 30-minute religion class), oversaw her students in twice-daily prayers, and took her class to the school-wide monthly Mass. The court found these religious duties minimal compared to Perich’s in Hosanna-Tabor. But it also noted that the teacher lacked a ministerial title or “credentials, training, or ministerial background” and did not hold herself out as a minister.

A California appeals court, in Su v. Stephen S. Wise Temple, similarly relied on Biel to hold that teachers at a Jewish day preschool were not ministers despite their function of “transmitting Jewish religion and practice to the next generation.” The teachers taught religion in the classroom, taught “Jewish rituals, values and holidays, [led] children in prayers, celebrate[d] Jewish holidays, and participate[d] in weekly Shabbat services.” But the court said they were non-ministers because of their secular title of “teacher,” their lack of “any formal Jewish education or training,” and the lack of a requirement that they be ordained or even Jewish.

The Supreme Court has granted review in Morrissey-Berru and Biel. In particular, neatly frames the issue left open by Hosanna-Tabor: Should an employee with unquestionably significant religious responsibilities fall outside the definition of minister because a court determines she lacks a title, training, ordination, or other ministerial credentials?

II. “IMPORTANT RELIGIOUS FUNCTION” VERSUS OTHER FACTORS

When an employee has important religious functions in a religious organization, there should be no further requirement that the employee have a ministerial title, training, background, or other credential. Focusing on important functions without imposing any further credentialing requirements suits the ministerial exception’s rationales, follows from the original understanding of the Religion Clauses and the evils to which it responded, and avoids denominational inequality and judicial intrusion into religious questions.

A. The Ministerial Exception’s Rationales

A function-based definition of minister best fits the chief rationales for the ministerial exception. The first of those rationales, as emphasized in Hosanna-Tabor, is to preserve religious organizations’ “control over the selection of those who

21 Id. at 208-09 (internal quotations marks and citations omitted).
22 882 F.3d 655, 661-62 (7th Cir. 2018).
23 Id. at 661.
24 Id.
25 700 F.3d 169, 177 (5th Cir. 2012).
26 Id.
27 Id. at 180.
28 769 Fed. Appx. at 461.
30 769 Fed. Appx. at 461.
31 Id.
32 911 F.3d 603.
33 Id. at 605.
34 Id. at 606-09. The court opined that if the mere presence of one Hosanna-Tabor circumstance was sufficient to make that employee a minister, the rest of the Supreme Court’s opinion would be rendered “irrelevant dicta.” Id. at 609. This is unpersuasive. Employees with significant religious functions should be considered ministers, regardless of their credentials. But when there is a question whether those functions are significant, courts may look to the title, training, and holding out of employees to determine whether to call them ministers. See infra section II.
35 Id. at 608-09.
37 Id.
will personify [their] beliefs”: those “who will preach their beliefs, teach their faith, and carry out their mission.”39 The Court concluded, “The church must be free to choose those who will guide it on its way.”40

Given that rationale, what should trigger the ministerial exception is the employee’s performance of religious functions, since such functions are what the organization needs to control. As Justice Alito’s concurrence put it, “Religious autonomy means that religious authorities must be free to determine who is qualified to serve in positions of substantial religious importance.”41 Whether the point sounds in free exercise (the church’s right to choose persons for such positions) or non-establishment (the ban on government choosing them), religious function is the touchstone.

A second rationale articulated for the ministerial exception is that adjudicating a minister’s discrimination suit will entangle a court in improperly deciding religious questions. This entanglement occurs because any discriminatory intent behind a firing, which an employer seldom expresses explicitly, must usually be proven circumstantially by showing that other proffered reasons were pretext. As Judge Richard Posner wrote, in applying the exception to dismiss a suit by a church’s organist/music director:

[T]he diocese would argue that [the plaintiff] was dismissed for a religious reason—his opinion concerning the suitability of particular music for Easter services. . . . Tomic would argue that the church’s criticism of his musical choices was a pretext for firing him, that the real reason was his age. The church would rebut with evidence of what the liturgically proper music is for an Easter Mass and Tomic might in turn dispute the church’s claim. The court would be asked to resolve a theological dispute.42

Justices Alito and Kagan made a similar point in their Hosanna-Tabor concurrence: if the parties dispute the employer’s intent, then “[i]n order to probe the real reason for [the plaintiff’s] firing, a civil court—and perhaps a jury—would be required to make a judgment about church doctrine.”43 If the religious reason the organization asserts for the firing were shown to be “an obscure and minor part of [its] doctrine, it would be much more plausible” for the plaintiff to argue it was a pretext for discrimination, while an asserted reason that is “a central and universally known tenet . . . would seem much more likely to be nonpretextual.”44 The civil court would be setting standards for the effectiveness of a minister, which in other contexts has been found to be clearly impermissible.45

The Hosanna-Tabor majority made clear that the ministerial exception does not stem solely from the Establishment Clause ban on judges inquiring into religious questions; it also stems from the substantive right of churches to choose their leaders under the Free Exercise Clause. The exception, the Court said, does not “safeguard a church’s decision to fire a minister only when it is made for a religious reason” but “instead ensures that the authority to select and control who will minister to the faithful . . . is the church’s alone.”46 Nevertheless, the bar on religious questions is one valid rationale for the exception.

The “no religious questions” rationale, like the substantive right to choose leaders, suggests that the definition of a minister should focus on the employee’s function. An employee’s religious functions are what make the inquiry into the employer’s motive improper and entangling: the court has to determine whether the organization’s complaints about the employee’s performance of such functions were weak enough to be pretextual. Again, the exception’s purposes are served by triggering it for jobs with significant religious functions.

In short, in Justice Alito’s words, the applicability of the ministerial exception “rests not on [the employee’s] ordination status or her formal title, but rather on her functional status as the type of employee that a church must be free to appoint or dismiss in order to exercise the religious liberty that the First Amendment guarantees.”47

B. Ministerial Credentialing and Original Meaning

The recent Ninth Circuit and California decisions narrow the definition of minister by requiring ministerial credentials as well as religious function. These decisions depart from the original meaning and understanding of the Religion Clauses. Such narrow definitions, especially through requirements of ministerial education or credentials, were among the evils that helped spur adoption of the First Amendment. As Hosanna-Tabor noted, religious establishments involved government appointment and control of ministers; it was “against this background that the First Amendment was adopted.”48 The founding era public would have understood government setting of credentials for ministers as a violation of the free exercise of religion and as an aspect of an establishment of religion.

The Constitution’s religious freedom guarantees arose in significant part from disputes between established colonial churches and pietistic evangelical dissenters, including both Baptists and New Light Congregationalists.49 From 1740 to 1754, the New Lights separated from the Old Light Congregationalist establishment, dissatisfied with its “‘formality’ [and] spiritual

[40] Id. at 196.
[41] Id. at 200 (Alito, J. concurring).
[42] Tomic v. Catholic Diocese of Peoria, 442 F.3d 1036, 1040 (7th Cir. 2006).
[44] Id.
[47] Id. at 206 (Alito, J., concurring).
[48] Id. at 182-83. See also, e.g., Payton v. New York, 445 U.S. 573, 584, 585-86 (1980) (interpreting the Fourth Amendment to apply to cases that raise “the evil the Amendment was designed to prevent” and “against which its wording . . . is directed”) (quotation omitted).
The New Lights, emerging from the revivals of the Great Awakening, spoke "to men's hearts or souls, to their spiritual emotions, not to their understanding or minds." Naturally, this attitude reflected who they chose to teach their faith. The New Lights opposed the formally trained "legal preacher," preferring a "layman who had experienced conversion personally." They loathed the "implication that since only an exceptionally intelligent and well-educated man could fathom the doctrinal mysteries of religion, the laws of nature, and the philosophy of science, salvation was only likely for the elite, the intelligentia." They believed that "the learned clergy had lost touch with the spiritual needs of the common man and no longer really served as ministers of God to them."54

Similar views about ministry arose among the so-called Separate Baptists, who likewise grew as a result of revivals to become a large dissenting group in both New England and the South. Although they differed from New Light Congregationalists by rejecting infant baptism, they likewise affirmed the work of itinerant, evangelistic preachers whose foundational authority was the divine call more than formal learning.55

New England colonial legislatures, which reflected the views of the Old Lights, responded by taking steps to restrict or disfavor informally trained ministers.56 In 1742, Connecticut passed a law prohibiting "itinerants" from preaching without the approval of an established parish. The same year, it passed legislation preventing any church or parish from choosing a minister who was not "educated at some university, college or publick [sic] academy" or who did not have "a degree from some university, college, or such public academy."57 The only alternative for a prospective pastor was to have "obtained testimonials" from the majority of "settled ministers of the gospel" in the county where he sought to minister finding him "to be of sufficient learning to qualify [sic] him for the work of such ministry."58

Similarly, Massachusetts passed a law in 1760 preventing legal recognition of parish ministers unless they had "academy or college training, or had obtained testimonials from the majority of the ministers already settled in the county." The law disqualified uncredentialed ministers, primarily Baptists, from receiving funds that were collected by each town's authorities for support of worship.60

The dissenters viewed these laws as religious freedom violations and religious establishments. Isaac Backus, a leader among Massachusetts Baptists, cited the colony’s law as an example of how the "blend[ing]" of "civil and ecclesiastical affairs . . . depriv[ed] many of God’s people of that liberty of conscience which he has given them."61 Backus argued that by "compel[ling] each parish to settle a minister" but then disqualifying teachers who lacked the government’s preferred training, the law clashed with the theological truth that God "gives gifts unto men in a sovereign way as seems good unto him."62 The law therefore forced dissenters to "render unto Caesar" something "that belongs only to God," since God "always claimed it as his sole prerogative to determine by his own laws what his worship shall be, who shall minister in it, and how they shall be supported."63 This and other aspects of the religious-tax system led the Massachusetts Baptists in 1773 to begin a "massive civil disobedience campaign" against it.64

In Virginia, civil authorities dictated where ministers were permitted to preach and jailed unlicensed ministers (most of whom were itinerant, non-establishment preachers). A young James Madison wrote to a friend, impassionedly, that such restrictions reflected a "diabolical, hell-conceived principle of persecution."65 The regulations stemmed from a deep-seated fear that the itinerants—who ignored parish boundaries, preached at times and places of their choosing, and disregarded the Book of Common Prayer—would "give great Encouragement to fall off from the established Church if they [were] permitted to range f[rom] the established Church if they [were] permitted to range and raise Contributions over the whole Country."66

Leaders among Virginia’s establishment, both civic and religious, worried about the itinerants, complaining of the "assemblies, especially of the common People, upon a pretended religious Account; convened sometimes by merely Lay

51 Id.
52 Id.
53 Id. at 352.
54 Id.
56 1 McLoughlin, supra note 50, at 363.
57 Id. at 472-73; see also id. at 363.
58 Id. at 473.
59 Jacob C. Meyer, Church and State in Massachusetts 1740–1833 51 (1930).
The newcomers were deemed dangerous precisely because they lacked the credentials that traditionally denoted ministers. In 1745, Governor William Gooch wrote against "false teachers . . . who, without order or license, or producing any testimonial of their education or sect, . . . lead the innocent and ignorant people into all kinds of delusion."69 Upon the arrival of a new contingent of preachers, Anglican clergyman Patrick Henry, Sr. (uncle of the statesman) declared, "I wish they could be prevented, or, at least be oblig'd to show their credentials."70

Disputes like these helped spur the adoption of the First Amendment. By the time the Constitution was being framed, religious dissenters in Virginia had secured freedom from many of these restrictions,71 and they feared a federal government capable of resurrecting the restrictions. Many New England dissenters had pushed for the new Constitution to provide greater protection than it ultimately did. John Leland, a Baptist minister in both Massachusetts and Virginia, complained that the unamended Constitution provided no "Constitutional defence" against religious oppression of the type Baptists had suffered.72 In 1787, most Baptists were antifederalist—opposing the Constitution's ratification—principally because of their dissatisfaction with its limited protections for religious freedom.73

James Madison owed his 1789 election to Congress to disgruntled Baptists who supported his candidacy in part to address their grievances with the established church in Virginia.74 Madison then made good on his promise to dissenters, introducing what became the Bill of Rights and taking a leading role in securing Congress's approval. He later reported that a Baptist leader assured him that the Bill of Rights "had entirely satisfied the disaffected of his sect."75

In short, narrow definitions of minister—notably, those setting educational and other credentials for ministers—were prominent among the evils to which the Religion Clauses were a response. Today, some courts are repeating this evil by effectively requiring that a minister possess "credential[s], training, or ministerial background" in order for an organization to invoke the ministerial exception.76 Such requirements impose civil authorities' assumptions—almost inevitably majoritarian assumptions—that certain training or formalities are inherent in the concept of a minister.

The Eighteenth-Century colonial laws used narrow definitions of minister to deny congregations their choice of preacher or teacher, or to deny ministers public funds that were available to those with training the government deemed adequate. Today, some courts use a similarly narrow definition to deny religious organizations the protection of the ministerial exception, exposing them to employee lawsuits that threaten the organizations' ability to choose who will teach the faith. The evil is the same in each case: subjecting religious organizations to a legal burden or disability regarding their chosen leaders based on those leaders' lack of credentials.

It is irrelevant that the colonial establishmentarians and today's judges may have different reasons for imposing these narrow, credential-based definitions. The colonial legislatures wanted to maintain social order and proper religion and worried that untrained ministers would "give great Encouragement to fall off from the established Church."77 The Ninth Circuit panels want to maintain maximum legal protection for employees by minimizing the scope of the ministerial exception. But the nature of the motivation does not matter to the ministerial exception, which protects religious autonomy even against laws that are generally applicable and have secular purposes.78 Whatever the motive, civil rules that require credentials for one to qualify as a minister perpetuate historic evils which the First Amendment aimed to prevent.

C. Denominational Inequality and Judicial Second-Guessing

Among the principles promoted by the Religion Clauses are religious non-discrimination and government neutrality on religious questions. Excluding an employee with important religious functions from minister status based on a lack of title, training, or other credentials creates a preference for some faiths over others and invites courts to second-guess religious organizations' self-understanding, in contravention of those important constitutional principles.

1. Inequality Among Faiths

"The clearest command of the Establishment Clause is that one religious denomination cannot be officially preferred over another."79 Requiring ministerial title, training, ordination, or other credentials as criteria for minister status invites

---

69 Id. (emphasis added).
70 Id.
71 "The legal status of toleration in Virginia remained uncertain until resolved by the revolutionary Declaration of Rights in 1776." Id. at 153.
73 1 McLoughlin, supra note 50, at 556-57; 1 Stokes, supra note 66, at 309. The Baptists were not a monolith concerning ratification. Backus, although generally representative of Baptist thought, went against the majority of his coreligionists in voting for ratification as a delegate to the Massachusetts ratifying convention, citing the exclusion of religious tests in Article VI as one of "our greatest securities in this constitution." Id. But believing that the test-oath ban was sufficient to justify ratifying the Constitution is perfectly consistent with believing, as he did, that laws other than test oaths also violated religious freedom and established religion.
75 McConnell, supra note 49, at 1487 (quoting letter from Madison to President Washington).
76 Morrissey-Berru, 769 Fed. Appx. at 461; Biel, 911 F.3d at 608.
77 See Lambert, supra note 67, at 127.
78 See supra note 3 and accompanying text.
79 Larson v. Valente, 456 U.S. 228, 244 (1982).
discrimination against small or minority faiths, religions with non-hierarchical polities, and faiths that use schools to sustain their beliefs. As Justices Alito and Kagan noted in their Hosanna-Tabor concurrence, “it would be a mistake if the term ‘minister’ or the concept of ordination were viewed as central to the important issue of religious autonomy.” Such criteria would disadvantage faiths that do “not employ the term ‘minister,’” that “eschew the concept of formal ordination,” or that (like Quakers, for example) “consider the ministry to consist of all or a very large percentage of their members.” “Because virtually every religion in the world is represented in the population of the United States,” broad application of the ministerial exception is necessary to protect minority religions.

Justice Thomas likewise warned that definitions of minister must be flexible and deferential because our nation includes religious organizations with “different leadership structures and doctrines that influence their conceptions of ministerial status,” and courts should avoid “disadvantaging those religious groups whose beliefs, practices, and membership are outside of the ‘mainstream’ or unpalatable to some.”

Conditioning the applicability of the ministerial exception on a religious employer’s use of certain terminology in job titles leads to unequal treatment of different faiths. The term minister itself can produce discrimination among religions, as it has strong Protestant associations and is rarely if ever used in this way by Catholics, Jews, Muslims, Hindus, or Buddhists.

The Ninth Circuit relied on the absence of formal ministerial titles in finding that religious school teachers were not ministers. These decisions confirm Justice Alito’s warning that such a focus will create improper inequalities among religious organizations. In Morrissey-Berru, the panel asserted that although the teacher had “significant religious responsibilities,” her “formal title of ‘Teacher’ was secular”; in Biel, the panel majority said that the “teacher” title did not “convey[ ] a religious—as opposed to secular—meaning.” That approach discriminates, at the very least, against religious groups that rely heavily on teachers and schools to “transmit[ ] the[ir] faith to the next generation.”

A number of faiths show such reliance: the Supreme Court has recognized that teachers commonly play a “critical and unique role . . . in fulfilling the mission of a church-operated school.”

As Judge D. Michael Fisher noted in his dissent in Biel, the formal title “Grade 5 Teacher” should be interpreted in the light of the employer’s “expression of [a teacher’s] role in the school,” which was the religious role of “a distinctively Catholic Grade 5 Teacher.” The fact that such schools call their employees by their most conventional, accurate title—teacher—does not detract from the fact that the employees perform a critical religious function. These schools should not be excluded from the protections of the ministerial exception because they choose that accurate title while others choose one that a court deems more minister-like.

Similar problems arise from a focus on whether the employee had official ministerial training. Throughout history—including at the time of the founding—some religious groups have strongly believed that God can anoint or call preachers, teachers, and leaders without formal religious education. Eighteenth-Century Baptists believed, in Isaac Backus’s words, that God “gives gifts unto men [e.g. preaching and teaching] in a sovereign way as seems good unto him.” Similarly, although Quakers have taken varying positions over time on ministers’ education, in their early years they “repudiated the idea that ministers must be scholastically trained. God called those who were to preach, and that was the only qualification necessary or possible.” Requiring ministerial training to qualify churches for the legal benefit of the ministerial exception discriminates against such groups.

In addition, a specialized training requirement, like a title requirement, discriminates against religions that rely on school teachers to communicate the faith to their students. In Morrissey-Berru, for example, the Ninth Circuit found the teacher’s “substantial religious responsibilities” insufficient because the “teacher” title did not reflect “ministerial substance and training.”

That suggests, improperly, that the religious training needs to be of the sort the court deems suitable for a clergy-like minister—a standard that most school teachers, even religiously important ones, will not meet.

Finally, a training requirement can also discriminate against small and minority religious groups. Such groups may

80 565 U.S. at 198 (Alito, J., concurring).
81 Id. at 202 (Alito, J., concurring). See Friends General Conference, FAQ: about Quakers, https://www.fgcquaker.org/discover/FAQs-about-quakers (“Quakers believe that we are all ministers and responsible for the care of our worship and community. Rather than employing a pastor, Quaker meetings function by appointing members to offices and committees.”).
82 Hosanna-Tabor, 565 U.S. at 198 (Alito, J., concurring); see also American Legion v. American Humanist Ass’n, 139 S. Ct. 2067, 2094 (2019) (Kagan, J., concurring) (emphasizing the importance of “sensitivity to and respect for this Nation’s pluralism”).
83 565 U.S. at 197 (Thomas, J., concurring) (citing Corporation of Presiding Bishop v. Amos, 483 U.S. 327, 336 (1987)).
84 Id. at 198 (Alito, J., concurring).
85 760 Fed. Appx. at 461; Biel, 911 F.3d at 608 (quotation omitted, brackets adjusted).
86 Hosanna-Tabor, 565 U.S. at 192.
88 911 F.3d at 616 (Fisher, J., dissenting) (emphasis added).
89 Backus, supra note 61, at 317 (italics removed).
90 Thomas D. Hamm, The Quakers in America 86 (2003). Early Quaker leader Robert Barclay gave a characteristic explanation:

We do believe and affirm, that some are more particularly called to the Work of the Ministry; and therefore are fitted of the Lord for that purpose. . . . That which we oppose, is, the distinction of Laity and Clergy . . . whereby none are admitted unto the work of the Ministry, but such as are Educated at Schools on purpose.

Id. at 86 (quotation omitted; italics in original).
91 Morrissey-Berru, 769 Fed. Appx. at 461 (emphasis added).
92 Moreover, the Ninth Circuit compounds its mistake by running title and training together: the panel majority in Biel reasoned, in criticizing the dissent, that for the teacher to be a minister, her title should
lack the resources to provide formal training programs or may lack sufficient candidates who have undergone such training. Teachers in these faiths may fail to qualify as ministers under the Ninth Circuit's analysis, even when performing the same religious function as teachers of other faiths with more resources for training.

2. Judicial Second-Guessing and Resolution of Religious Questions

Requiring ministerial title or training invites another First Amendment evil: it requires courts to resolve questions of religious doctrine, second-guessing an organization's own determinations about what features are most important in constituting leadership roles in the organization. As the Supreme Court held in *Serbian Eastern Orthodox Diocese v. Milivojevich*, courts must accept the decisions of ecclesiastical tribunals regarding their own rules and regulations for internal discipline and government.93 Judicial second-guessing of these ecclesiastical decisions is an impermissible substitution of the church's internal governance.94 *Milivojevich* forbade courts from second-guessing a church's decision to discipline and defrock of one of its bishops; it held that the decision to fire or discipline a minister was a "quintessentially religious" controversy.95 *Hosanna-Tabor* relied on these principles in concluding that "it is impermissible for the government to contradict a church's determination of who can act as its ministers."96

But the right to choose or discipline ministers "would be hollow . . . if secular courts could second-guess the organization's sincere determination that a given employee is a 'minister' under the organization's theological tenets."97 Accordingly, a broad, flexible definition of "minister" is necessary to avoid resolving essentially religious controversies. "[C]ivil courts are in no position to second-guess a religious organization's decision as to whether an employee's 'religious function . . . made it essential that she abide by [the employer's] doctrine' and decision-making."98

Excluding employees with important religious functions from the category of ministers on the ground that they lack adequate titles or training would bring on precisely these evils. It would require courts to determine what sort of title is sufficiently minister-like to qualify. Again, the cases excluding teachers are instructive. Contrary to the Ninth Circuit's holdings,99 there is nothing inherently secular about the title "teacher." It can communicate the important religious function of teaching religious doctrine and values—especially, as already noted, when the title is used in the setting of a school that is grounded in and teaches a religious faith.100 Yet under the Ninth Circuit's approach, religious schools seeking the shield of the ministerial exception—freedom to select leaders without judicial second-guessing—would be forced to rechristen their employees with titles more aligned with what the court considers to be religious.

Similar problems arise with requiring ministerial training. In entrusting important religious functions to various employees, including teachers, an organization typically prescribes the training it believes necessary or appropriate for those functions. Yet the Ninth Circuit's recent approach holds that an employee who has not received training that the court deems suitable for a minister is not a minister. Under this approach, courts must decide just what sort and extent of training is enough. A more entangling inquiry could hardly be imagined. Thus, as the Seventh Circuit recently observed, the Ninth Circuit's approach improperly embraces "independent judicial resolution of ecclesiastical issues."101

An approach that requires credentials on top of an employee's significant religious functions would allow courts to "second-guess" a religious organization's assessment that an employee's "religious function . . . made it essential that [she] abide by [the employer's] doctrine" and decision-making.102 In the face of civil liability, some religious groups may be pressured to change their practices: spending additional resources on clergy-like training, relying on ordained persons rather than laymen to teach the faith, or shifting their religious teaching away from K-12 classrooms. The Supreme Court has warned that when judges can second-guess organizations on such religious matters in civil lawsuits, "[f]ear of potential liability might affect the way an organization carried out what it understood to be its religious mission."103 Organizations will be pressured, as Justice Thomas has put it, to "conform [their] beliefs and practices regarding 'ministers' to the prevailing secular understanding."104

D. Holding Out as a Minister

The final consideration mentioned in *Hosanna-Tabor* is whether the employee was "held out" as a minister, either by

---

94 Id. at 708; see also Presbyterian Church v. Hull Church, 393 U.S. 440, 449 (1969) ("First Amendment values are plainly jeopardized when church property litigation is made to turn on the resolution by civil courts of controversies over religious doctrine and practice.").
95 426 U.S. at 720.
96 *Hosanna-Tabor*, 565 U.S. at 185.
97 Id. at 197 (Thomas, J., concurring).
98 Id. at 206 (Alito, J., concurring).
99 See *supra* note 85 and accompanying text.
100 See *supra* notes 87-88 and accompanying text.
101 Sterlinski v. Catholic Bishop of Chicago, 934 F.3d 568, 570 (7th Cir. 2019) (Easterbrook, J.) (criticizing *Biel* for "essentially disregarding what *Biel*'s employer . . . thought about its own organization and operations" and for largely ignoring "whether the employee served a religious function").
102 *Hosanna-Tabor*, 565 U.S. at 206 (Alito, J., concurring); see also id. at 197 (Thomas, J., concurring).
103 *Amos*, 483 U.S. at 335 ("[I]t is a significant burden on a religious organization to require it . . . to predict which of its activities a secular court will consider religious.").
104 *Hosanna-Tabor*, 565 U.S. at 197 (Thomas, J., concurring).
the employee herself or by the religious organization. This factor can raise complications. It can be the vehicle for a proper focus on employees' religious functions: courts can legitimately require that the employer communicate those functions, that is, hold out the employee as performing them. But holding out can also prompt the errors of requiring credentials of employees and ignoring their religious functions.

As an example of the first error, the Ninth Circuit in Biel ruled that the school did not “hold [the teacher] out as a minister by suggesting to its community that she had special expertise in Church doctrine, values, or pedagogy.” By focusing on expertise, the court required the employer to communicate the employee's ministerial training and credentials, not merely her substantial religious function. This is simply another way of requiring such credentials, and thus it suffers from the flaws with credentialing detailed above.

Even if the court properly interprets holding out to communicate function rather than credentials, it can still err unless it focuses on the employer's understanding of function, not just the employee's unilateral view. To focus on the employee's unilateral action of holding out or not invites the civil court to resolve ecclesiastical disputes. In every ministerial exception case where the definition of minister is at issue, the plaintiff claims an understanding of the term different from the organization's understanding and invites the court to impose the employee's understanding on the organization through civil liability. In other words, the plaintiff asks the court to engage in the very second-guessing—the very resolution of ecclesiastical questions—that the Supreme Court has said is improper.

III. The Boundaries of Important Functions

For the reasons above, employees who perform important or significant religious functions should be deemed ministers regardless of whether they also have credentials such as a ministerial title, training, education, or ordination. Those other features should not be irrelevant: when it is a close question whether an employee's religious functions are significant, the employee's status might still be determined with reference to title, training, or other credentials. But the presence of significant religious functions should be sufficient to establish that an employee is a minister, even if it is not necessary in every case.

The focus on functions does not mean that all employees of religious organizations will qualify as ministers. The exception needs to have boundaries, especially because within its bounds it is absolute. The ministerial exception cannot be overridden by a compelling governmental interest; it therefore applies to all claims of discrimination, even discrimination based on race. Moreover, unlike many other free exercise exemptions, the ministerial exception is not limited to cases where discrimination is motivated by religious doctrine, such as the male-only rule for

---

105 Id. at 191.
106 See Biel, 911 F.3d at 608.

---

108 Hosanna-Tabor, 565 U.S. at 194-95 ("The purpose of the exception is not to safeguard a church's decision to fire a minister only when it is made for a religious reason."); Simpson v. Wells-Lamont Corp., 494 F.2d 490, 493 (5th Cir. 1974) (rejecting rule limiting the exception "to differences in church doctrine").
109 Hosanna-Tabor, 565 U.S. at 204 (Alito, J., concurring); see id. at 202-04 (citing lower court decisions); supra notes 17-27 and accompanying text.
111 Hosanna-Tabor, 565 U.S. at 193-94.
look to them as leaders). Of course, this need not mean leading a congregation, or an entire organization. Individual teachers are ministers because they lead their audience—students—in learning religion or engaging in prayers or worship.

Nor will all teachers in religious schools qualify as ministers under a functional definition. However, the following activities, at least, indicate minister status: (1) the teacher teaches a class in religion, with some inculcation of religious principles; (2) the teacher is tasked with integrating religion into other subjects taught; or (3) the teacher engages or supervises students in religious observances such as chapel, prayers, Bible readings, or special religious programs. There should be evidence that the teacher not only is assigned such duties (for example, by a school handbook\textsuperscript{112}) but also actually carries out the duties.

IV. Conclusion

An employee’s significant religious functions should be sufficient to make the employee a minister for the purposes of the ministerial exception. In close cases, courts should also look to employees’ title and training, but the absence of such credentials should never trump the presence of significant religious function. This function-focused inquiry avoids the evil of state-sponsored ministerial credentialism, a practice that helped motivate the adoption of the First Amendment. Focusing on function also furthers the fundamental Religion Clause principles of equality among denominations and judicial non-involvement in the ecclesiastical decision-making of religious organizations. The Supreme Court should call a halt to the recent trend of credentialism in some lower courts, which threatens to undermine the purposes of the ministerial exception.

\textsuperscript{112} See, e.g., the handbook description of religious duties in \textit{Morrissey-Berru}, 760 Fed. Appx. at 461; \textit{supra} notes 29-30 and accompanying text.
**Kisor v. Wilkie Makes Auer a Paper Tiger**

By Karen R. Harned & Aeron Van Scoyk

**Administrative Law & Regulation Practice Group**

**About the Authors:**

Karen R. Harned serves as Executive Director of the NFIB Small Business Legal Center, a 501(c)(3) public interest law firm that serves as the voice for small business in the nation’s courts and a legal resource for small business owners nationwide. Aeron Van Scoyk is a Law Clerk with the NFIB Small Business Legal Center and a student at U.C. Davis School of Law. Any opinions set forth in this article are attributable to the authors and not the organization they represent.

**Note from the Editor:**

The Federalist Society takes no positions on particular legal and public policy matters. Any expressions of opinion are those of the authors. Whenever we publish an article that advocates for a particular position, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. We also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.

**Other Views:**

allowed benefits to flow starting in 1983. The denial was based on a determination that the combat records were not “relevant” under (c)(1)’s process for a reconsideration of the denied claim because the materials didn’t establish PTSD as a current disability and, as a result, would not have changed the result at the time. According to the VA, records had to be “outcome determinative” to be found relevant under (c)(1). Kisor argued that a record should be deemed relevant if it has “any tendency to make the termination of the action more [or less] probable.”

The Federal Circuit agreed with the VA. It acknowledged that the term “relevant” was ambiguous in the regulations and applied Auer deference. It quoted Auer saying that “[a]n agency’s interpretation of its own regulations is controlling unless plainly erroneous or inconsistent with the regulations being interpreted.” Deferring to the VA’s interpretation, it held that Mr. Kisor’s benefits should be calculated under the less generous “reopened” provision.

The Supreme Court vacated the decision below. In an opinion authored by Justice Elena Kagan, the Court upheld Auer deference, but not before seriously circumscribing its applicability. Justice Kagan argued that reflexive application of Auer deference is a “caricature of the doctrine.” Before accepting an agency’s interpretation of a regulation, a court must perform its own independent analysis and exercise the courts’ traditional “reviewing and restraining functions.” A court must exhaust all of its traditional tools of construction before concluding that a rule is ambiguous. Even if reading an administrative rule makes “the eyes glaze over,” a court should not “wave the ambiguity flag” until it has thoroughly analyzed the regulation. In other words, courts may no longer rubber stamp an agency’s interpretation of its regulation and cite to Auer.

Even if an administrative rule is “genuinely ambiguous,” the agency’s interpretation of that rule must still be “reasonable.” Rebuffing Seminole Rock’s “plainly erroneous” formulation, Justice Kagan cabin’d the reasonableness inquiry by focusing on three important markers. First, Justice Kagan noted that the regulatory interpretation must be the agency’s “authoritative” or “official position.” An off-hand comment by an agency employee or an informal memo is not enough to state an agency’s interpretation of a regulation.

More importantly, the agency’s interpretation must reflect its area of substantive expertise. If the presumption justifying Auer deference is that Congress delegates its lawmaker authority to an agency because of its administrative knowledge and expertise, then it makes no sense to defer to an agency’s interpretation of a topic outside of that scope. Although the majority opinion did not weigh in on the specific interpretation at issue in Kisor, it noted that certain provisions may be better interpreted by a judge, such as the meaning of a common law property term or a question of the award of attorney’s fees. An evidentiary issue might very well be more in a judge’s wheelhouse than that of an administrative agency.

Finally, an agency’s interpretation is rarely reasonable if it is inconsistent. A post hoc interpretation that serves as a “convenient litigating position” does not warrant Auer deference. A court may not defer to an agency’s interpretation that creates “unfair surprise” or upsets settled reliance interests. For example, imposing retroactive liability for longstanding conduct that had never before been addressed by the agency does not warrant Auer deference.

II. Concurring Opinions Focus on the Limits of Auer Deference

Justice Gorsuch concurred in the decision, but he would have jettisoned Auer entirely. He thought the majority’s new limitations left Auer deference on “life support,” and that the Court would have to resolve remaining issues at a later date. Tracing the history of Auer deference back to its roots in Seminole Rock, Justice Gorsuch opined that the doctrine was never intended to be more than dicta. The “controlling weight” discussion in Seminole Rock was not central to the holding in that case, yet increasingly the Supreme Court and lower courts “mechanically applied and reflexively treated” that dicta.

Justice Gorsuch argued that courts faced with cases that turn on the interpretations of regulations should be applying the Administrative Procedure Act (APA), not “writing on a blank slate or exercising some common-law-making power.” And the APA’s directives are clear. Section 706 of the APA directs courts to “decide all relevant questions of law” and “set aside agency action . . . found to be . . . not in accordance with law.” Further, courts must “determine the meaning” of any “agency action.” Thus courts must decide for themselves the best meaning of a disputed

---

10 Id. at 1366.
11 Id. at 1367.
12 Kisor, 139 S. Ct. at 2415.
13 Id.
14 Id.
15 Id.
16 Id. at 2416.
17 Id.
18 Id. at 2417.
19 Id.
20 Id.
21 Id. at 2418.
22 Id.
23 Id. at 2425.
24 Id. at 2428-29.
25 Id. at 2429.
26 Id. at 2432.
28 Id.; see 5 U.S.C. 551(13) (defining “agency action”).
Gorsuch’s approach is likely to be similar, if not the same.40 That the practical result of either the majority approach or Justice Kagan’s reasoning under Article III.37 Judges should be able to independently analyze a regulation, with the agency’s interpretation providing persuasive, but not controlling, weight so long as the agency offers a valid rationale for a consistent interpretation within its area of expertise.38

Moreover, Justice Gorsuch voiced his concern that Auer deference is at odds with the Constitution’s separation of powers to the extent that it prevents judges from exercising their duty under Article III.39 Judges should be able to independently analyze a regulation, with the agency’s interpretation providing persuasive, but not controlling, weight so long as the agency offers a valid rationale for a consistent interpretation within its area of expertise.38

Both Chief Justice John Roberts and Justice Kavanaugh wrote separate concurrences that emphasized a similar point: there is not much distance between Justice Kagan and Justice Gorsuch’s opinions. Chief Justice Roberts noted that the majority’s prerequisites for applying Auer deference are very similar to Justice Gorsuch’s list of reasons a court might find an agency’s interpretation influential.39 Justice Kavanaugh wrote to emphasize that the practical result of either the majority approach or Justice Gorsuch’s approach is likely to be similar, if not the same.40 While formally rejecting Auer would have been more direct, both approaches require a judge to appropriately scrutinize a regulation by employing the traditional tools of construction, which will “almost always” lead to the best interpretation of the issue, negating the need to defer to an agency’s interpretation at all.41

III. A Philosophical Split on the Role of Stare Decisis

The opinions reveal a split in judicial philosophies about the role of stare decisis, “the special care [the Justices] take to preserve [their] precedents” but that “is not an inexorable command.”42 Five of the Justices—Justices Kagan, Ginsburg, Breyer, and Sotomayor in the controlling opinion and Chief Justice Roberts in his concurrence—agreed that stare decisis weighed against overturning Auer. Overruling Auer, Justice Kagan reasoned, would mean overturning “dozens of cases” the Supreme Court has decided, as well as “thousands” of lower court decisions.43 Overturning Auer would upset many settled constructions of rules, particularly in the context of administrative law. These Justices also seemed to be wary of opening the floodgates for fresh challenges to settled administrative rules; Justice Kagan quoted the Solicitor General’s assessment at oral argument that every ruling based on Seminole Rock would need to be relitigated.44 Moreover, Justice Kagan reasoned, Congress could step in at any time to amend the APA, lessening the need for the Supreme Court to act as a final backstop and overrule Auer.45

Justices Gorsuch, Thomas, and Kavanaugh did not share the majority’s concerns regarding stare decisis.46 Rather, they pointed to the explosive growth of the administrative state as a reason for not following stare decisis. Justice Gorsuch contended that Auer has not generated serious reliance interests because an agency’s expectation of Auer deference as an entitlement is not a legitimate interest when weighed against “the countervailing interest of all citizens in having their constitutional rights fully protected.”47 Instead, the number of cases decided under Auer only magnifies the harm that could be corrected by overturning Auer. While reliance interests have long been a factor the Court considers in its stare decisis analysis, Justice Gorsuch indicated that not all interests should carry the same weight. Rather, interests closer to the core of the Constitution, such as “the interests of citizens in a fair hearing before an independent judge,” are more important than “the convenience of government officials.”48 Justice Gorsuch also noted that the majority’s imposition of new limitations on Auer would lead to many settled decisions being relitigated anyways.49

---

29 Kisor, 139 S. Ct. at 2432.
30 Id.
31 Id. at 2434; 5 U.S.C. § 553.
33 Kisor, 139 S. Ct. 2400 at 2434.
34 Id.
35 Id.
36 Id. at 2444-45.
37 Id. at 2437.
38 Id. at 2447.
39 Id. at 2424-25.
40 Id. at 2448.
41 Id.
42 Id. at 2418 (Kagan, J.); Id. at 2445 (Gorsuch, J., concurring).
43 Id. at 2422.
44 Id.
45 Id. at 2422-23.
46 Justice Alito expressed no opinion on the role of stare decisis in this case.
47 Id. at 2447.
48 Id.
49 Id.
IV. Early Post-Kisor Decisions Show Lower Courts Are Cutting Back on Auer Deference

Lower courts have begun applying the new Kisor standard in a variety of contexts. For instance, the Ninth Circuit considered a dispute between the Secretary of Labor and OSHA over conflicting interpretations of respirator use regulations. The Ninth Circuit sided with the secretary, concluding that the regulation at issue was unambiguous and refusing to defer to OSHA’s interpretations under Auer. The same court considered Amazon’s challenge to the IRS’s interpretation of its regulations that resulted in reallocating income from a European subsidiary back to the U.S., finding that the IRS had incorrectly characterized Amazon’s European assets and refusing to accord Auer deference.

In at least one high-profile case, a court has struck down an agency’s interpretation using the factors outlined in Kisor. In Romero v. Barr, the Fourth Circuit considered whether the U.S. Attorney General’s interpretation of the Immigration and Nationality Act warranted Auer deference. In 2018, the Attorney General issued a decision determining that immigration judges and the Board of Immigration Appeals do not have authority under existing regulations to administratively close an immigration proceeding. Applying the recently announced factors from Kisor, the Fourth Circuit held that the Attorney General’s interpretation failed at every step and therefore did not merit Auer deference.

First, the court, applying traditional tools of interpretation, concluded that the regulations at issue “unambiguously provide [immigration judges] and the [Board of Immigration Appeals] with the general authority to administratively close cases.” Even if the regulation were somehow ambiguous, the Attorney General’s reading would still not warrant deference because it amounts to “unfair surprise”—the new interpretation “breaks with decades of the agency’s use and acceptance of administrative closure” and does not give “fair warning” to the regulated parties. After going through the Kisor factors, the Fourth Circuit finally noted that the Attorney General’s reading was not persuasive under Justice Gorsuch’s test either because it “comes too late in the game.”

V. Next Stop—Chevron Deference?

Finally, the Kisor decision serves as a preview to a much-anticipated showdown over Auer’s big brother, Chevron. For at least two decades, there has been a fierce debate within the administrative law bar about whether the Chevron doctrine has outlived its usefulness. The doctrine is named after Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc., in which the Supreme Court said that when a statute is ambiguous, judges are to defer to the interpretation of the federal agency charged with implementing it, as long as that interpretation is reasonable. Proponents of Chevron argue, among other things, that agency experts are better positioned than judges to understand the practical implications of the statutes they are implementing via regulations. When discerning the requirements of an ambiguous statute, a judge should not override the agency expert’s reasonable interpretation. Instead, she should defer.

Chief Justice Roberts specifically noted in his concurrence that his opinion in Kisor does not “touch upon” the issue of Chevron deference. Justices Kavanaugh and Alito expressly joined this sentiment. Yet at least five Justices seem to be skeptical of, or at least willing to reexamine, Chevron. Justice Gorsuch in his concurrence, which Justice Thomas joined, wrote in a footnote that “there are serious questions . . . about whether [Chevron deference] comports with the APA and the Constitution.” On the other side, Justice Kagan’s opinion in Kisor approvingly cited to Chevron, perhaps to reaffirm the continuing vitality of the doctrine. Justices Breyer, Ginsburg, and Sotomayor all joined Justice Kagan’s opinion in full. It remains to be seen where all the Justices will ultimately come down.

We may get some answers to this term if the Supreme Court decides to take Baldwin v. United States. Baldwin addresses Brand X deference, a subset of Chevron deference. Under Brand X, an agency’s interpretation of a statute is due Chevron deference regardless of whether it is inconsistent with prior practice. The question in Baldwin is whether an administrative agency can overrule a court’s prior interpretation of a statute. In this case, the Ninth Circuit had previously construed a statute regulating the postmark date of tax documents filed with the IRS, an important issue when there is a dispute about whether a tax document was timely filed. The Ninth Circuit held that the statute did not displace the common law mailbox rule. The mailbox rule allows proof of mailing to establish a presumption that a document is physically delivered on the date of the postmark. In 2011, the IRS amended its regulations to hold that the statute does displace the common law mailbox rule, disallowing taxpayers from presenting

51 Id. at 1310.
52 Amazon.com, Inc. v. Comm’r of Internal Revenue, 934 F.3d 976 (9th Cir. 2019).
53 937 F.3d 282 (4th Cir. 2019).
54 Id. at 286.
55 Id. at 292.
56 Id. at 295.
57 Id. at 297.
60 Kisor, 139 S. Ct. at 2425.
61 Id. at 2446, n.114.
63 Petition for Certiorari at 3–5, Baldwin v. United States, No. 19-402 (Sept. 23, 2019).
64 Id. at 7–8.
evidence to establish a postmark date.\textsuperscript{65} Under \textit{Brand X}, the IRS's new interpretation is given \textit{Chevron} deference even though it conflicts with a prior court ruling.\textsuperscript{66}

VI. Conclusion

In the few months since \textit{Kisor v. Wilkie} was decided, it appears that \textit{Auer} 2.0—the newly limited version of the doctrine formulated in Justice Kagan’s opinion—may end up being the paper tiger Justices Roberts and Kavanaugh predicted it would be. It is becoming clear that the Justices recognize that a day of reckoning is coming—sooner rather than later—for the judicial doctrines that have aided and abetted an ever-growing administrative state.

\footnotesize{\textsuperscript{65} Id. at 8-9. \textsuperscript{66} Id. at 9.}