Towards an Administrative Rule of Lenity: Restoring the Constitutional Congress by Reforming Statutory Interpretation
— Joel S. Nolette (winner of the Article I Initiative Writing Contest)

The Problem with the Proliferation of Collateral Consequences
— John G. Malcolm

What Happened to the Public’s Interest in Patent Law?
— Kristen Jakobsen Osenga

Compelled Speech in *Masterpiece Cakeshop*: What the Supreme Court’s June 2018 Decisions Tell Us About the Unresolved Questions
— James A. Campbell

Impeachment: The Constitution’s Fiduciary Meaning of “High . . . Misdemeanors”
— Robert G. Natelson

The GDPR: What It Really Does and How the U.S. Can Chart a Better Course
— Roslyn Layton & Julian McElendon

Party Like It’s 1935?: *Gundy v. United States* and the Future of the Non-Delegation Doctrine
— Matthew Cavedon & Jonathan Skrmetti

A Shy Frog, the Administrative State, and Judicial Review of Agency Decision-Making: A Preview of *Weyerhaeuser v. United States Fish & Wildlife Service*
— Mark Miller
Letter from the Editor

The FEDERALIST SOCIETY REVIEW is the legal journal produced by the Federalist Society’s Practice Groups. The REVIEW was formerly known as ENGAGE, and although the name has changed, it still features top-notch 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 and full issues are available at 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
Independent Review of Procurements Is Worth It: There Is No Support for Hamstringing the GAO Bid Protest Process

By Marcia G. Madsen, David F. Dowd, Roger V. Abbott

Note from the Editor:
This article criticizes a recent change to the GAO bid protest process. A new rule requires bidders whose protests in large procurements fail to pay DoD’s costs of processing the protest. The article argues that this cost-shifting provision is out of step with the APA’s goals of promoting effectiveness and integrity in agency actions, including government contracting.

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.


About the Author:

Marcia Madsen is a partner at Mayer Brown, where she chairs the Government Contracts practice and co-chairs the National Security practice. David Dowd is also a partner at Mayer Brown, whose litigation practice has a strong emphasis in government contracting issues and controversies. Roger Abbott is a Litigation & Dispute Resolution associate at Mayer Brown.

Over the past two years, critics from the Department of Defense (DoD) and Congress have claimed that frivolous or unnecessary bid protests are impairing the procurement process, especially the ability of DoD to obtain weapons systems and services in a timely manner. In response to these complaints, the Senate Armed Services Committee (SASC) considered changes to the Government Accountability Office (GAO) bid protest process. One of the changes the SASC considered was to penalize contractors that file unsuccessful bid protests at GAO involving large defense procurements by requiring them to pay DoD’s costs of processing the protests. This loser pays proposal was included in the National Defense Authorization Act (NDAA) for Fiscal Year (FY) 2018 as Section 827.1 The provision is an unwarranted effort to undermine independent review of agency procurement actions.

This cost-shifting provision reflects a basic failure to appreciate the importance of independent review of government procurement decisions. Contract awards are agency decisions that involve billions of dollars in taxpayer funds. The new loser pays provision—an English-style cost-shifting rule—violates basic principles of administrative law enshrined in the Administrative Procedure Act (APA), which was designed to protect against arbitrary, capricious, and illegal government action. Notably, efforts by private sector defendants to impose a similar cost-shifting approach have been largely rejected, even though they are not subject to the same constitutional restrictions as government agencies.2 This type of rule penalizes citizens for attempting to vindicate their rights by seeking review of government decisions, which no other agency review process does. As explained below, the government already enjoys a deferential standard of review in bid protests. Shifting the costs of litigation to unsuccessful protesters sends a very clear message to contractors: DoD’s largest procurements are not for review. This message is inconsistent with the right of citizens to seek independent review of government actions.

Additionally, the asserted basis for restricting review is without factual support. At the time Section 827 was proposed and enacted, there was relatively little data on bid protests. The existing data was largely limited to the statistics published by GAO and the Court of Federal Claims (CFC). These reports did not provide granular information on issues of concern, such as the

---

1 Pub. L. No. 115-91 (Sec. 827).
2 Since the late 18th century, the United States has rejected the loser pays “English Rule” and generally requires each party to bear its own litigation expenses (an approach known as the “American Rule”). Although a number of exceptions to this rule have emerged since the turn of the 20th century, these exceptions have been narrowly tailored to shift costs in favor of successful plaintiffs rather than the defendant, as here. David A. Root, Attorney Fee-Shifting in America: Comparing, Contrasting, and Combining the “American Rule” and “English Rule”, 15 IND. INT’L & COMP. L. REV. 583, 584–89 (2005) (noting that fee-shifting is available for successful plaintiffs in four categories of cases: civil rights suits, consumer protection suits, employment suits, and environmental protection suits).
The number of procurement actions (including task orders) versus the number of protests filed, or the prevalence of bid protests by incumbent contractors. There was no data supporting the notion that protests of large acquisitions are hampering procurement efforts, and certainly not to an extent that would justify restricting normal rights of citizens to seek review of government action. The RAND Corporation was tasked by Congress with developing data for a study. The RAND report, which was issued to Congress on December 21, 2017, refutes the notion that protests are a problem. Significantly, RAND was not asked to review whether changing the bid protest process would restrict citizens’ right to petition for review of government action.

Finally, and curiously, limiting or reducing review of major defense procurement decisions is incompatible with DoD’s stated aim of improving competition and eliminating corrupt agency behavior. As it stands, fewer than 50% of DoD acquisitions are competitively sourced. Any change that discourages independent review of agency procurement decisions will impair the government’s ability to promote competition and minimize corruption.

I. Bid Protests Provide an Important Vehicle, Firmly Rooted in the APA, for Ensuring That Agencies Act Lawfully

Although discussions about possible changes and reforms tend to focus very heavily on public contract laws and regulations, the award and administration of government contracts is—in practice—agency decisionmaking involving billions of dollars in taxpayer funds. Administrative law principles are therefore an important consideration in the regulation of agency procurements. Agency decisions of all types, including government contracting, are broadly governed by the APA, which “creates the framework for regulating executive agencies” by, among other things, providing for independent review of agency decisions to counterbalance the power of large government agencies like DoD.

A. The APA Sets the Framework for Review of the Exercise of Power by Government Agencies

The APA created the framework for regulating the modern administrative state. The APA was enacted in 1946 in response to the expansion and centralization of federal power under the New Deal, which had resulted in the proliferation of enormously powerful administrative agencies. As one scholar put it, the APA “established the fundamental relationship between regulatory agencies and those whom they regulate. . . . The balance that the APA struck between promoting individuals’ rights and maintaining agencies’ policymaking flexibility has continued in force, with only minor modifications, until the present.”

Following the flurry of legislation in 1933, President Roosevelt’s New Deal coalition eventually began to falter. Concerned by the dangers posed by the rapid centralization of power as exhibited in Germany, members of Congress launched a campaign for administrative reform. This effort culminated in the Walter-Logan administrative reform bill, which was passed by Congress and vetoed by Roosevelt in 1940. The reforms proposed in Walter-Logan were much more restrictive than those in the later APA. Among other things, Walter-Logan included much more thoroughgoing provisions for judicial review than the APA (including a very broad standard for standing), required notice and public hearings (rather than just notice and comment) for all new rules or rule changes, and even required that agencies enact any regulations pursuant to their enabling statutes within one year of the passage of those statutes.

Even with Roosevelt’s veto of Walter-Logan, reform efforts continued unabated and became even more active following Roosevelt’s death. Both parties eventually settled on the APA as a compromise measure that would, over time, protect the advances made by the regulatory state while giving citizens and businesses tools to check the arbitrary exercise of power by agencies. The

---

3 The threshold for protesting DoD task orders is currently $25 million. 10 U.S.C. § 2304(e)(1)(B). The threshold for protesting civilian task orders, or DoD task orders issued by a civilian agency, is $10 million. 41 U.S.C. § 4106(f)(3). HP Enterprise Services, LLC, B-413382.2, Nov. 30, 2016, 2016 CPD • 343 (holding that the jurisdictional threshold for civilian task orders applies to DoD task orders issued by civilian agencies, like the GSA).

4 NDAA for FY 2017, Conference Report to accompany S. 2943, sec. 885.


6 Defense Procurement Acquisition Policy publishes quarterly competition scorecards regarding DoD acquisitions. In the first and second quarters of FY 2017, the percentage of procurements that were sourced competitively was 47% and 49%, respectively. See http://www.acq.osd.mil/dap/opic/competition.html.


9 In a 1937 message to Congress, President Roosevelt noted that “[t]here are over 100 separate departments, boards, commissions, corporations, authorities, agencies, and activities through which the work of the Government is being carried on.” Franklin D. Roosevelt, Message from the President of the United States (Jan. 12, 1937) in THE PRESIDENT’S COMMITTEE ON ADMINISTRATIVE MANAGEMENT, REPORT OF THE PRESIDENT’S COMMITTEE ON ADMINISTRATIVE MANAGEMENT iii-iv (1937).


11 Id. at 1581-87.


13 See Shepherd, supra note 10 at 1598-1601.
balance struck by this hard-fought compromise is reflected in the bid protest process.

B. The Bid Protest Process Provides a Check on the Administrative State

According to the Attorney General’s Manual on the Administrative Procedure Act, the APA seeks to balance the requirements of due process and sound administration in the following four ways: (1) by “requir[ing] agencies to keep the public currently informed of their organization, procedures, and rules”; (2) by “provid[ing] for public participation in the rulemaking process”; (3) by “prescrib[ing] uniform standards for the conduct of formal rulemaking and adjudicatory proceedings”; and (4) by defining the scope of judicial review in the context of the administrative state.14

Of particular relevance here, the APA confers a broad right of judicial review to parties directly affected by agency conduct. Pursuant to Section 10 of the APA, “A person suffering legal wrong because of agency action, or adversely affected or aggrieved by agency action within the meaning of a relevant statute, is entitled to judicial review thereof.”15 The precise scope of judicial review is found in Section 706 of the Act. Among other things, Section 706 authorizes the courts to decide questions of law and “hold unlawful and set aside agency action, findings, and conclusions found to be . . . arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.”16 In Motor Vehicle Manufacturers Association, the Supreme Court clarified that:

[A]n agency rule would be arbitrary and capricious if the agency has relied on factors which Congress has not intended it to consider, entirely failed to consider an important aspect of the problem, offered an explanation for its decision that runs counter to the evidence before the agency, or is so implausible that it could not be ascribed to a difference in view or the product of agency expertise.17

In 1970, in the landmark Scanwell Laboratories, Inc. v. Shaffer case,18 the D.C. Circuit held that the protections afforded by the APA against arbitrary action by agencies apply to agency procurements. The court explained that the APA “embodies the basic presumption of judicial review to one suffering legal wrong because of agency action” and held that Section 10 confers standing on disappointed offerors to sue the agency in federal court.19 Although the D.C. Circuit acknowledged that “the ultimate grant of a contract must be left to the discretion of a government agency,” the court held that it is:

[I]ncontestable that that discretion may not be abused. . . . Contracting officers may not base decisions on arbitrary or capricious abuses of discretion . . . and our holding here is that one who makes a prima facie showing alleging such action on the part of an agency or contracting officer has standing to sue under section 10 of the [APA].20

The D.C. Circuit made plain that arbitrary and capricious action by an agency includes violating the terms of the solicitation and failing to comply with procurement laws and regulations.21 The loser pays provision of the 2018 NDAA is inconsistent with the APA’s presumption of judicial review for a citizen who suffers legal wrong because of agency action. There is no basis for penalizing citizens for trying to challenge arbitrary and capricious agency conduct.

Although the standard of review applied by GAO in evaluating agency conduct is not governed by the APA or defined by statute or regulation, the GAO applies the same Scanwell standard in its approach to review. In an advisory opinion made at the request of the federal district court for the District of Columbia, GAO noted that its “standard of review comports with the [D.C. Circuit’s] standard that provides deference to the decisions of procurement officials; an agency’s procurement decision will only be disturbed where it involves ‘a clear and prejudicial violation of applicable statutes or regulations’ or ‘had no rational basis.’”22

II. GAO’s Bid Protest Process is Rooted in APA Concepts

A. The Evolution of GAO into an Effective Bid Protest Forum

Although the bid protest process at all tribunals is rooted in concepts underlying the APA, this section focuses on the GAO, as the GAO in particular has been the target of recent reform efforts. The GAO has been an active administrative forum for bid protests for nearly 100 years. Due to its informal and expeditious process, GAO handles a large number of protests every year without resort to the courts.23 The GAO, headed by the Comptroller General of the United States, began as the General Accounting Office, and was established through the Budget and Accounting Act of 1921.24 Among other things, GAO considered whether public

18 424 F.2d 859 (D.C. Cir. 1970).
19 Id. at 869.
20 Id. at 869. ("When the bounds of discretion give way to the stricter boundaries of law, administrative discretion gives way to judicial review").
21 Florida Prof’l Review Org., Inc.—Advisory Opinion, B-253908, Jan. 10, 1994, 94-1 CPD ¶ 17 n. 20.
22 Florida Prof’l Review Org., Inc.—Advisory Opinion, B-253908, Jan. 10, 1994, 94-1 CPD ¶ 17 n. 20.
23 According to GAO’s annual bid protest reports to Congress, 2,433 bid protests were filed in FY 2017, 2,586 in FY 2016, 2,496 in FY 2015, 2,445 in FY 2014, 2,298 in FY 2013, and 2,339 in FY 2012, for an average of 2,433 protests filed per year. NB: these figures are slightly lower than the number of “cases filed” for each year, as they exclude cost claims and requests for reconsideration. These reports are available at https://www.gao.gov/legal/bid-protest-annual-reports/about.
funds spent by agencies had been appropriated by Congress, and it was directed to report to Congress “every . . . contract made by any department in any year in violation of law.”

Several years after its formation, GAO began to consider bid protests from disappointed bidders as an adjunct of its authority to settle and adjust claims by and against the United States. GAO was the only forum for bid protests until 1956, when the Court of Claims asserted jurisdiction to hear such protests, based on the theory that the Tucker Act granted disappointed offerors standing to claim damages when the government violated its implicit contractural duty to evaluate bids in good faith. But GAO’s effectiveness as a bid protest forum was severely compromised by its inability to grant enforceable relief. Until the introduction of the automatic stay in 1984, agencies “frequently responded to the filing of a bid protest, or other form of Congressional concern over how certain resources were being purchased, by rushing to award a contract and begin its execution.” As a result, “most procurements became fait accompli before they could be reviewed.” Once awarded, even a contract that was the product of a material failure to comply with legal requirements was a done deal.

To remedy this “major loophole,” when Congress enacted the Competition in Contracting Act of 1984 (CICA), it enhanced the effectiveness of GAO by providing a short and temporary automatic stay of a contract award and a suspension of ongoing performance during the pendency of the protest if the protest was timely filed with GAO (time periods for filing are short and strictly enforced). Pursuant to CICA, “a contract may not be awarded in any procurement after the Federal agency [conducting the procurement] has received notice of a protest with respect to such procurement from the Comptroller General and while the protest is pending.” An agency that believes it cannot wait the 100 days can override the stay by making a written finding that “urgent and compelling circumstances which significantly affect the interest of the United States will not permit awaiting the decision of the GAO” and reporting this finding to GAO. The Court of Federal Claims now has exclusive jurisdiction over CICA override challenges. The number of overrides historically is small.

Although the GAO’s recommendations are not binding on agencies, unlike judgments made by the Court of Federal Claims, as a practical matter, agencies almost always follow GAO recommendations. For instance, in both FY 2016 and FY 2017, the agencies uniformly followed all GAO recommendations, and in FY 2015, only one GAO recommendation was disregarded by an agency. As a result, most bid protests are resolved without resort to the courts.

B. Complaints about GAO Bid Protests and Proposed Changes

Notwithstanding the availability of stay overrides and the fact that GAO is required to resolve bid protests within 100 days, critics from DoD have expressed concerns about the state of the bid protest system. They argue, among other things, that major procurements are routinely bottled up by “frivolous protests.” These critiques are not new.

Although the SASC considered several changes to the bid protest process to address these critiques during its markup in 2016 of the NDAA for FY 2017, these proposals were ultimately rejected by the Conference Committee in favor of a proposal to have “an independent research entity . . . with appropriate

---

33 1 U.S.C. § 3553(c)(1); see also FAR 33.104(b)(1).
34 1 U.S.C. § 3553(c)(2)-(3); see also FAR 33.104(b)(1).
39 For example, in FY 2017, the GAO resolved 2,433 bid protests. In contrast, the Court of Federal Claims resolved 133 bid protests.
40 See, e.g., Kovacic, supra note 27 at 489–91.
41 SASC inserted a so-called “loser pays” provision, which would have “require[d] a large contractor filing a bid protest on a defense contract with GAO to cover the cost of processing the protest if all of the elements in the protest are denied in an opinion issued by GAO.” Report of the SASC on the NDAA for FY 2017 at Title VIII Sec. 821. The SASC also included other provisions, such as a measure to discourage incumbent protests. Id.
expertise” perform a “report on bid protests.”42 This report, which was presented to Congress on Dec. 21, 2017, was to include:

- An analysis of “the extent and manner in which the bid protest system affects or is perceived to affect” various aspects of the procurement process, including the use of discussions and decision to use sole source award methods;
- An examination of bid protest trends, including the number of protests filed in each forum, the overall ratio of protests to procurements, and the overall effectiveness of protests at different forums; and
- “[A]n analysis of bid protests filed by incumbent contractors” inquiring into all sorts of factors, including the rate at which such protests are filed, the delay caused by these protests, how often these protests are sustained, and how often protesters are ultimately awarded the contract that is subject to the protest.

Notwithstanding the impending report on bid protests, Congress proceeded to include a variant of the loser pays provision in the NDAA for FY 2018. Section 827 requires DoD to establish a pilot program within two years of passage of the bill, to “require[] contractors to reimburse [DoD] for costs incurred in processing covered protests,” which include those filed by companies with revenue in excess of $250 million that are denied by GAO.43

As discussed below, the newly released RAND study does not provide any evidence supporting the loser pays provision. Regardless, this measure contradicts the fundamental concept behind the APA—that “a person suffering legal wrong because of agency action, or adversely affected or aggrieved by agency action within the meaning of a relevant statute, is entitled to review.”44

There are more reasons the measure is irrational. For one thing, disappointed bidders do not even have access to any part of the administrative record until 30 days after the protest has been filed;45 they must decide whether to protest based on the information provided by the agency in the debriefing (if there is one) or notice of award, which is very limited. Furthermore, the deadline to file a bid protest is stringent, which decreases the likelihood of frivolous protests; in order to avail themselves of the automatic stay, disappointed bidders must file a protest within five calendar days after a required debriefing, if there is one, or within 10 days after the date of contract award.46 Finally, the fact that GAO declines to sustain a protest does not establish that it was unreasonable to file the protest. Agencies are entitled under the APA to substantial deference in review of their actions. The reasonableness of agency action can only be examined once the record is produced. Just because a protester ultimately cannot overcome the deferential standard does not mean that the allegations lacked merit. Indeed, the substantial “effectiveness” rate at GAO (approximately 45% of protests are either sustained or subject to agency corrective action prior to decision) demonstrates that there is substantial merit perceived by agencies in many cases. Furthermore, cases that are denied on a written opinion frequently reflect a close call on the merits.

In short, discouraging bid protests is contrary to the basic principle behind the APA—that independent review of agency decisionmaking is necessary to counterbalance the accretion of power by administrative agencies. Additionally, as explained below in Part III.C, this provision frustrates DoD’s own stated objectives of encouraging competition and preventing corruption in government contracting.

III. The Purpose and Benefits of a Meaningful Review of Agency Procurement Actions

A. A Critical Oversight Role: Protests Help Ensure that Agencies Act in Accordance with the Law

1. Public Contracting is Fundamentally Different from Commercial Contracting

Government contracting is different in many fundamental respects from commercial contracting. Since government contracts are financed using funds from the public fisc, government contracts are highly regulated; in addition to the Federal Acquisition Regulation (FAR), there are numerous agency FAR supplements. Contractors are subject to a number of government-unique enforcement statutes and regulations, including the Truth in Negotiations Act,47 the False Claims Act,48 various anti-kickback49 and anti-bribery statutes,50 domestic preference statutes such as the Buy American Act,51 and various

---

42 NDAA for FY 2018, Pub. L. 115-91 (Sec. 827).
43 NDAA for FY 2017, Pub. L. 114–328 (Sec. 885).
45 31 U.S.C. § 3553(b)(2) (setting a 30-day deadline for a normal, non-expedited protest, and a 20-day deadline for an expedited protest).
47 The Truth in Negotiations Act requires certain contractors (in negotiated, or non-commercial, procurement actions exceeding $750,000) to disclose “cost or pricing data,” certify the data is accurate, complete, and current, and lower their prices to reflect any price increase caused by a defective disclosure. 10 U.S.C. § 2306a. This requirement applies to all contracts that are priced or performed on the basis of cost.
48 Pursuant to the False Claims Act, any “person” who “knowingly presents, or causes to be presented” a “false or fraudulent claim” to the U.S. Government is liable for treble damages and civil penalties. 31 U.S.C. § 3729(a).
50 For instance, federal law prohibits any person, such as a contractor, from directly or indirectly giving, offering, or promising anything of value to agency officials for or because of any official act performed or to be performed by such official. 18 U.S.C. § 201(c)(1)(a).
51 The Buy American Act requires that the U.S. Government purchase only “manufactured articles, materials, and supplies” that “have been mined or produced in the United States” and such “manufactured articles, materials, and supplies” that “have been manufactured in the United States substantially all” from U.S. components, unless doing so is “inconsistent with the public interest” or would result in “unreasonable” cost. 41 U.S.C. §§ 8301-8305 (formerly codified at 41 U.S.C. §§ 10a – d).
regulations prohibiting the use of foreign counterfeit parts, human trafficking, and more. Because they spend taxpayer funds, agencies face a number of restrictions as buyers that do not affect buyers in the commercial sector. First and foremost, authority for government contracts must be provided by Congress in order to be lawful. Additionally, agencies must conduct procurements using taxpayer funds in accordance with a number of laws, such as the Competition in Contracting Act, which require agencies to open up procurements to competition and deal fairly with offerors. More broadly, agencies are under a general obligation to conduct procurements in a reasonable manner and to avoid acting arbitrarily and capriciously. This fundamental requirement is found not only in procurement laws but also in the APA, which broadly governs the conduct of agencies. This is in marked contrast to commercial buyers, who are not spending taxpayer money and can make sourcing decisions unconstrained by regulation—they can make non-competitive contracts for reasons other than the merit of the product or service offered, for example. The only limits to such private behavior are set by the market.

2. GAO Protests Provide Effective and Efficient Oversight

GAO bid protests effectively subject agencies to scrutiny by exposing their decisionmaking (as reflected in the agency record) to real time review within 100 calendar days. Although relatively few procurements are actually protested, the possibility of a protest encourages agency officials to act lawfully and provides a remedy for unlawful conduct. Protesters, as private attorneys general, are better situated to know the circumstances of procurements in which they participate than other sources of after-the-fact oversight, such as agency inspectors general or prosecutors.

Contrary to assertions that too many protests are filed, in FY 2012 through 2017, an average of 2433 bid protests were filed each year at GAO—one protest for every $192 million in procurement spending. Against that backdrop, GAO sustains a relatively small number of protests each year. For instance, in FY 2017, GAO sustained 99 protests, or 17% of all GAO decisions made on the merits. “[T]he most prevalent reasons for sustaining protests during the 2017 fiscal year were: (1) unreasonable technical evaluation; (2) unreasonable past performance evaluation; (3) unreasonable cost or price evaluation; (4) inadequate documentation of the record; and (5) flawed selection decision.” This data is consistent with the findings of the RAND study, which are discussed below.

Additionally, as GAO points out, numerous protests that are not sustained are nonetheless “effective” because they spur agencies to review the matter internally and to take corrective action before GAO issues an opinion. Thus, in addition to publishing a “sustain” rate, GAO includes an “effectiveness rate” in its reports to Congress, which describes the frequency with which the protester receives “some form of relief from the agency, as reported to GAO, either as a result of voluntary agency corrective action or [GAO] sustaining the protest.” Notably, the effectiveness rate has largely held constant over the past ten years despite the increasing number of protests. The effectiveness rate for FY 2017 was 47%.

GAO sustains protests due to some flaw in the evaluation process. But aside from encouraging agency compliance, GAO protests occasionally help forestall potentially catastrophic mistakes by agencies. For instance, in PCCP Constructors, the U.S. Army Corps of Engineers issued a solicitation for the design-build of permanent canal closures and pumps along three outfall canals at or near Lake Pontchartrain, Louisiana. In the event of severe flooding, such as that induced by Hurricane Katrina, the pumps help move water out of New Orleans and into the outfall canals, where the excess water can be diverted. In the protest that followed the initial award, the protester argued, among other things, that the pumping stations outlined in the awardee’s designs were unable to withstand pressure from flood water, and that the agency had failed to detect these defects because it had accepted the awardee’s blanket statements at face value without properly scrutinizing the technical proposals, as required by the RFP. The chair of the technical evaluation team conceded that, despite the potential for catastrophe, his team considered an important aspect of the awardee’s technical design for less than five minutes. GAO found that the agency had failed to meaningfully evaluate the awardee’s technical proposal and sustained the protest on this ground. GAO also sustained the protester’s organizational conflict of interest protest ground, finding that the agency had failed to consider the impact of the awardee’s hiring the agency’s 

---

52 DoD procurement regulations require government contractors to obtain electronic parts from the original manufacturer or an authorized aftermarket manufacturer, if possible. The rule requires contractors to vet contractor-approved suppliers and to “assume[] responsibility for the authenticity of the parts provided” by them. See DFARS Case 2014-D005 (codified in part in DFARS 252.246–708 Sources of Electronic Parts).

53 Government contractors must also comply with federal legislation and regulations designed to combat human trafficking. See FAR Subpart 22.17 and FAR 52.222-50.

54 For instance, CICA generally requires federal agencies to engage in “full and open competition” to obtain needed supplies and services—i.e., the Government must provide all interested parties the opportunity to submit a bid or proposal for the contract. 10 U.S.C. § 2302(3) (covering DoD procurements); 41 U.S.C. § 107 (covering civilian procurements); see also FAR 6.003.

55 See infra section III.A.2.

56 As explained above in footnote 23, this average includes only bid protests, and excludes both cost claims and requests for reconsideration.

57 The most recent fiscal year for which bid protest data is available is FY 2017. See supra note 23. The total outlay of government contracts in each fiscal year is available on USAspending.gov. In FY 2012 through 2017, the average annual outlay on government contracts was $467,989,537,132. See https://www.usaspending.gov/Pages/TextView.aspx?data=OverviewOfAwardsByFiscalYearTextView.


59 Id. at 2.


61 PCCP Constructors, JV; Bechtel Infrastructure Corp., B-405036, Aug. 4, 2011, 2011 CPD ¶ 156.

62 Id.
Chief of Program Execution of the Hurricane Protection Office, who had supervised the direction of this very procurement until his retirement from government service.

The facts in this particular case are striking. The solicitation for this $700 million procurement was developed over the course of two years, and the agency's failure to consider a critical feature of the awardee's technical proposal for more than five minutes could have had disastrous consequences for the city of New Orleans. But for the bid protests in this case—which were resolved within a fraction of the two years it took the agency just to issue the solicitation, a fatal flaw in the awardee's proposal might never have been discovered. This case also highlights the potential danger posed by organizational conflicts of interest, which can impair the objectivity of the evaluation team or give an offeror with inside knowledge or agency connections an unfair competitive advantage.

B. Concerns about Frivolous Protests or Abusive Litigation by Incumbents Are Unfounded

Critics of the GAO bid protest system point out that the number of nominal protests has steadily increased since 2007, with the number of bid protests filed rising from 1,411 in FY 2007 to 2,353 in FY 2011 to 2,538 in FY 2013 to 2,734 in FY 2016. Notably, the nominal number of protests filed in FY 2017 dropped to 2,596, the lowest number since FY 2013. In any event, these raw numbers can be misleading, as GAO's docketing process counts every supplemental submission rather than each protested procurement. GAO's statistics represent the total number of docket numbers ("B" numbers), not actual protests. Daniel I. Gordon, former Administrator for Federal Procurement Policy, estimates that approximately 1.6 docket numbers are assigned per protested procurement. Moreover, this nominal increase obscures the fact that only a miniscule percentage of government procurements are protested. Gordon estimates that in recent years, well over 200,000 contracts and protestable task orders are awarded each year. Of these, approximately 99.3% and 99.5% of procurements are not protested. GAO's statistics represent the total number of docket numbers ("B" numbers), not actual protests.

RAND noted that the sustain rate for Task Order protests is higher than it is for other procurements, suggesting those procurements are at greater risk for competition violations.)

Admittedly, the largest DoD contracts are protested at a higher rate: "the higher the dollar value, the greater the likelihood of a protest. For a company that loses the competition for a $100 million contract, with all the bid and proposal costs that competing entails, the additional cost of filing a protest may seem minimal," particularly if the loss of the contract is not clearly on the merits. Nonetheless, the overall picture is that bid protests are very rare. And it is reasonable that extremely large acquisitions should be subject to a higher degree of scrutiny.

These statistics are consistent with the findings of the RAND report. The study concluded that "bid protests are exceedingly uncommon for DoD procurements—less than 0.3% of DoD procurements are protested. RAND also undermined the notion that incumbent protesters file meritless protests in order to profit from bridge contracts. Although RAND found that incumbent protesters were slightly more likely to protest an award than non-incumbents, it also noted that the effectiveness rate of protests filed by incumbents is at least as high as those filed by non-incumbent contractors. In fact, RAND found that incumbent protests of task order awards have a significantly higher effectiveness rate than those filed by non-incumbents. For instance, the overall effectiveness rate for FYs 2015-2016 was 45.5% for all procurements, 47% in the case of non-incumbents protesting task orders, and 71% for incumbents protesting task orders. Additionally, RAND directly refuted the notion that large defense contractors are disproportionately slowing down the procurement process by filing meritless protests. This concern was the basis for the Section 827 pilot program, which only focuses on large defense contractors. To the contrary, RAND's data shows that "the largest 11 [government contracting] firms have remained relatively constant and may be slightly declining." What is more, "[t]he top 11 firms have higher effectiveness and sustained rates than the rest of the sample [though these rates are declining over time]—suggesting that they are possibly more selective in the protests they file and spend more resources developing their cases." Rather, RAND suggests that the rise in bid protests is driven by small businesses. RAND found it "striking" that 58% of procurement protests were filed by small businesses, which in FY 2016, cumulatively comprised only 15% of DoD contract dollars. RAND suggests that the increasing incidence of protests

63 See supra note 23.
64 For instance, requests for reconsideration and requests by successful protestors for reimbursement of costs are all given separate case numbers. After eliminating such supplemental filings, the number of cases filed is 1,276 for FY 2007, 2,214 for FY 2011, 2,298 for FY 2013, 2,586 for FY 2016, and 2,433 for FY 2017. See id.
66 Id. at 495.
67 The NDAA for FY 2008 amended the Federal Acquisition Streamlining Act to grant GAO jurisdiction to hear protests concerning task or delivery orders valued at more than $10 million. See National Defense Authorization Act for Fiscal Year 2008, Pub. L. No. 110-181, § 843, 122 Stat. 3, 237-39 (codified as amended at 41 U.S.C. § 4106(f)). Although the jurisdictional grant over civilian task orders was initially limited to three years, it was extended in 2011 and eventually made permanent under the GAO Civilian Task and Delivery Order Protest Authority Act of 2016. The current thresholds for protests are $10 million for civilian task orders and $25 million for DoD task orders. See supra note 3.
68 Gordon, supra note 65, at 497.
70 Id. at 64-65.
71 Id. at 33-34. These 11 firms cumulatively comprise nearly 42% of total obligated DoD dollars in FY 2016. Id. at 33.
72 Id. at 36.
By small business can be addressed by improving post-award debriefings. 73

RAND also performed a qualitative analysis of the perspectives of bid protest stakeholders and reported that “the perspectives of the bid protest system from DoD personnel and the private sector varied greatly.” On the one hand, DoD personnel expressed “general dissatisfaction with the current bid protest system”; on the other hand, private-sector representatives “strongly supported the bid protest system.” 74 Although the armed services “reported that they did not track or collect data on whether companies are more or less likely to file a bid protest as a normal course of their business strategy,” 75 they all expressed concern that contractors who lose follow-on awards are much more likely to protest a procurement than non-incumbents, that contractors file too many “weak” protests, and that “contractors have an unfair advantage in the contracting process by impeding timely awards with bid protests.” 76 In light of the absence of any data supporting these concerns, this apparent hostility to the bid protest process appears to merely reflect opposition to subjecting agency procurement decisions to independent review.

Not only are protests rare, the delay they cause is minimized by the statutory requirement that GAO resolve protests within 100 days. 77 Moreover, GAO “consistently close[s] more than half of all [DOD] protests within 30 days.” 78 Although critics point out that procurements can be further delayed if the GAO denies a protest and the protester files again in the CFC, this does not happen often. 79

Some have speculated that incumbents file frivolous protests in order to continue working during the pendency of the CICA automatic stay. This was one of the primary concerns motivating recent efforts by the Senate Armed Services Committee. However, in the 2009 GAO Report, GAO noted that the last protest described by GAO as frivolous was filed in 1996. 80 Additionally, GAO emphasized that it is authorized to dismiss frivolous protests sua sponte. 81

Given the paucity of bid protests relative to the total number of DoD or federal procurements, the dollar value and importance of many DoD procurements, and the fact that DoD agencies often take several years just to design and implement solicitations for major defense acquisitions, it seems doubtful that the 100-day maximum for resolving a GAO bid protest is too long to wait. And even on the rare occasion in which time really is of the essence, the agency always retains the authority to override the automatic stay. 82

C. Bid Protests Benefit the Government and the Public by Enhancing Competition and Protecting the Integrity of Public Procurements

1. Bid Protests Enhance Competition in Public Contracts

Competitive procurement, whether formally advertised or negotiated, is beneficial to the government. First, competition in contracting saves money. . . In addition to potential cost savings, agencies have been able to promote significant innovative and technical changes through negotiated competition for contract awards. . . Lastly, and possibly the most important benefit of competition, is its inherent appeal of “fair play.” Competition maintains integrity in the expenditure of public funds by ensuring that government contracts are awarded on the basis of merit rather than favoritism. 83

As explained in a report by DoD’s Office of Procurement Policy, “[t]he premise that underlies this strong preference for ‘full and open competition’ is the economic premise that has long been recognized by the courts as the basis for a free market economic system—that competition brings consumers the widest variety of choices and the lowest possible prices.” 84 The federal government has long recognized the benefits of competition. The principle that bid protests in a one year period. GAO also dismissed the last protest for abuse of process. See Latvian Connection LLC, B-413442, Aug. 18, 2016, 2016 CPD ¶ 194. Nonetheless, GAO did not describe the protest as “frivolous.” As GAO explained, the word “frivolous,” in the judicial context, has a very narrow, technical meaning, which does not apply here. See U.S. Gov’t Accountability Office, B-401197, Report to Congress on Bid Protests Involving Defense Procurements 10 (2009).

81 Id. at 11-12.
82 For FY 2002, the last year in which GAO included information on overrides in its annual report on protests, GAO reported that with respect to the 1,101 protests filed that year, the agency used its override authority on only 65 occasions. U.S. Gov’t Accountability Office, GAO-03-427R, Bid Protest Annual to Congress for FY 2002, at 3, 4 (2003).
agencies should manage procurements competitively whenever possible is enshrined in both the Armed Services Procurement Act of 1947 and the Federal Property and Administrative Services Act of 1949, which still govern the conduct of military and civilian procurements, respectively.

Nonetheless, by the late 1970s, many in Congress and the policy community became concerned that existing procurement statutes were inadequate in promoting competition. Among other things, the authors of the CICA conference report in 1984 stated that “a strong enforcement mechanism is necessary to insure that the mandate for competition is enforced and that vendors wrongly excluded from competing for government contracts receive equitable relief.” To promote competition, CICA made a number of significant changes to procurement law. CICA established “full and open competition” as the standard for federal procurements, prohibited non-competitive negotiations except in narrowly-defined circumstances, added specific planning and publication requirements, and—of particular relevance here—empowered GAO to act as a forum for bid protests. For the first time, GAO was explicitly granted independent statutory authority to hear protests. In so doing, Congress formally recognized the importance of GAO bid protests in promoting competition and enhanced GAO’s ability to act as an effective arbiter through the limited 100-day automatic stay provision. Since the passage of CICA, Congress has looked periodically at changing CICA’s mandate for “full and open competition,” and has declined to do so.

In August 2014, DoD published Guidelines for Creating and Maintaining a Competitive Environment and again promoted the benefits of “[c]ompetition [a]s the most valuable means we have to motivate industry to deliver effective and efficient solutions for the [DoD]. When we create and maintain a competitive environment, we are able to spur innovation, improve quality and performance, and lower costs for the supplies and services we acquire.”

DoD has also recognized the need to run procurements more competitively. For example, in 2014, DoD issued guidelines aimed at maximizing direct and indirect competition in its contract solicitation and awards process. The creation of the guidelines followed DoD’s recognition that it had been experiencing a declining competition rate and had not met its competition goals during the previous four years. In addition, Under Secretary of Defense Frank Kendall’s Better Buying Power 3.0 memorandum advocates removing barriers to commercial technology utilization, noting in part that “the Department can do a much more effective job of accessing and employing commercial technologies.” Notwithstanding these efforts, DoD Competition Scorecards indicate that fewer than 50% of DoD acquisitions are competitively sourced. The push to discourage bid protests is inconsistent with DoD’s policy of encouraging competition and innovation in contracting.

2. Bid Protests Help Preserve the Integrity of the Procurement Process

The U.S. government is the single largest buyer in the world. In FY 2017, federal agencies spent $383 billion on a wide range of goods and services to meet their mission needs. Given the vast amount of money at stake, the risks posed by potential corruption are very real. Notwithstanding the “presumption of regularity” that the courts apply when scrutinizing the conduct of agency contracting officials, it is hardly surprising that the government in general—and DoD in particular—has been beset by numerous procurement scandals. Bid protest scrutiny helps preserve the integrity of the competitive process. “Competition curbs fraud by creating opportunities to re-assess sources of goods and services reinforcing the public trust and confidence in the transparency of the Defense Acquisition System.”

Examples of corruption in federal public contracting abound. For instance, the FBI’s Operation III Wind, conducted between
1986 and 1988, was one of the largest procurement fraud investigations in U.S. history, and resulted in the prosecution of over 60 contractors, consultants, and government officials. The investigation ensnared a number of high-ranking officials, including an Assistant Secretary of the Navy, who was found to have accepted hundreds of thousands of dollars in bribes and illegal gratuities. According to the FBI website, this scandal “so shocked the nation that just five months after the case became public, new rules governing federal procurement were put into place,” including the Procurement Integrity Act of 1988.

In June 2003, a colonel who had been the commander of the U.S. Army’s Contracting Command Korea, a position in which he oversaw the approval of more than $300 million in contracts a year, was sentenced to four and a half years of prison for accepting $900,000 in bribes from South Korean construction companies. According to the original indictment, the defendant rigged $150 million worth of military service contracts in South Korea.

Furthermore:

One witness testified that he provided the men with prostitutes at a government conference in the U.S. Virgin Islands in exchange for contracts worth $1.4 million, according to a report published on Morningstar.com, a news and information service on financial markets. Another witness also said he provided prostitutes for the men and paid for a 1997 trip to Las Vegas to see a heavyweight title fight between Mike Tyson and Evander Holyfield.

In August 2004, the former chief of Plans, Requirements and Acquisitions for the Defense Information Systems Agency, was indicted for conspiracy to defraud the United States, receiving illegal gratuities, wire fraud, money laundering, conflict of interest, conspiracy to conceal records, obstruction of justice, and suborning perjury.

One of the more notable scandals in recent years centers around Darlene Druyun, who served for years as the senior career civilian procurement officer for the U.S. Air Force (USAF), second only to the Assistant Secretary of the Air Force for Acquisition. Druyun developed a reputation for her take-no-prisoners, risk-taking approach to managing a series of complex, multi-billion-dollar deals, and was known within the industry as “The Dragon Lady.”

During her tenure, Druyun pushed a number of initiatives to discourage disappointed offerors from seeking independent review of Air Force award decisions. For instance, on April 23, 1999, the Air Force announced a “Lightning Bolt” acquisition reform initiative that required all major USAF programs to have a program-level alternative dispute resolution (ADR) mechanism. To that end, the USAF signed corporate agreements with more than 40 of the largest defense contractors, which required them to use ADR. While ADR is widely used and helpful in many circumstances, the result with respect to protests was to shield agency decisions from outside review. Similarly, on April 17, 2001, USAF announced that internal guidance would be amended to include “issue identification and resolution” as a criterion for evaluating contractor past performance under the Contractor Performance Assessment Reporting System to incentivize use of ADR. This change would allow contracting officers to downgrade the past performance ratings of contractors that failed—in the agency’s view—to act proactively in identifying and resolving disputes.

This effort was eventually blocked by both the General Services Administration and the Office of Federal Procurement Policy. Although the Air Force ADR program is well regarded, this particular effort was problematic as it limited the availability of independent, outside review of agency decisions.

In the late 1990s, Druyun began negotiating a deal to lease one hundred KC767A tankers from Boeing. Druyun’s retirement from government service in November 2002, and her new $250,000 a year job as vice president of Boeing, created a firestorm of controversy. The ensuing investigation eventually revealed that while Druyun was negotiating several contracts with Boeing in her capacity as a senior procurement official, she was simultaneously negotiating jobs at Boeing for herself, her daughter, and her daughter’s fiancé, in violation of federal

---


104 Id.


---
conflict of interest laws. Although she initially denied that this conflict of interest influenced her actions as procurement official, she eventually admitted that the conflict did influence her judgment on several procurements, including the KC-767A tanker lease, the Small Diameter Bomb procurement, and a contract dispute over the C-17 H22 procurement.

Shortly after Druyun made these admissions in her criminal proceedings, other offerors for the contract to modernize the C-130 filed bid protests before the GAO. The protesters alleged that the proposals had not been evaluated in a fair and unbiased manner, and that the agency had violated conflict of interest laws. Although filed over three years after the award had been made, the GAO treated the protests as timely on the ground that the protesters had no reason to know the information disclosed in Druyun’s admissions. GAO sustained the protests:

[W]here, as here, the record establishes that a procurement official was biased in favor of one offeror, and was a significant participant in agency activities that culminated in the decisions forming the basis for protest, we believe that the need to maintain the integrity of the procurement process requires that we sustain the protest unless there is compelling evidence that the protester was not prejudiced.

A subsequent protest against the Air Force’s award to Boeing in connection to the development of the small diameter bomb was also sustained on similar grounds.

Finally, the U.S. Navy is currently embroiled in the “Fat Leonard” corruption investigation, which has been described as “the worst corruption scandal in Navy history.” In January 2015, defense contractor Leonard Glenn Francis pled guilty to bribery and fraud charges, and agreed to forfeit $35 million to the government. Francis bribed officers of the Seventh Fleet with prostitutes, money, and vacation in exchange for being allowed to overcharge the Navy for fuel, tugboats, barges, sewage removal, and other services, as well as food and water.

The ensuing investigation was recently expanded to include over 60 admirals and hundreds of other U.S. naval officers.

In short, given the vast amounts of money being spent on federal contracting, and the fact that these outlays are likely to increase over time, oversight of agency officials will continue to be an ongoing challenge for the government. Any reform that discourages the independent review of agency decisions increases the prospect that contracts awarded in such circumstances will not be addressed and will continue undisturbed.

IV. Conclusion

Today, the U.S. Government is the largest buyer in the world and is controlled, not by market forces, but by an enormous bureaucracy with thousands of employees. Accountability is at the heart of the post-war compromise that resulted in the APA, and it is essential for the legitimacy of the modern administrative state. The current bid protest system subjects agency actions involving billions of dollars to real time review—with an expeditious process that resolves disputes within 100 days. Because of the short stay during the protest, it is possible to timely correct abuses and errors. Given that significant procurements take years to germinate to the point of a solicitation and award, agency complaints about delays attributable to protests are simply not credible. Without protests, any review would take place years after the fact under an IG or through a False Claims Act lawsuit, and no remedy could undo the damage already done by an irrational or illegal procurement decision. Moreover, companies looking at entering the federal market would hesitate when they realize that they have no real remedy if they are disadvantaged by insider deals or otherwise flawed procurements.

The recently enacted change to the bid protest process should be repealed, as it discourages disappointed offerors from filing bid protests and exercising their right to challenge the actions of large bureaucracies. No evidence suggests that such a constraint on review of agency action will be beneficial—indeed, there are no facts at all that suggest bid protests are being abused. If implemented, the loser pays provision will severely compromise the competitive procurement process. This provision will discourage competition and hinder the effectiveness of the bid protest mechanism in promoting integrity and fairness in contracting. Many commercial companies already regard federal procurement as “inside baseball,” which should concern proponents of attracting more commercial technology companies to the federal marketplace. Undermining meaningful review of agency actions is a step in the wrong direction for the future of federal contracting.

112 Id.
118 See Craig Whitlock, “Fat Leonard” Probe expands to ensnare more than 60 admirals, supra note 116.
Towards an Administrative Rule of Lenity: Restoring the Constitutional Congress by Reforming Statutory Interpretation

By Joel S. Nolette

Note from the Editor:
In this article, Joel Nolette proposes an “administrative rule of lenity” as a solution to structural problems besetting our federal government and the constitutional separation of powers.

Mr. Nolette’s article won First Place in the Article I Initiative Writing Contest, conducted by the Federalist Society’s Article I Initiative on the topic Restoring the Constitutional Congress. Prof. Lillian BeVier, Hon. C. Boyden Gray, and Hon. Chris DeMuth were the esteemed judges for the Writing Contest. They completed a blind review of the submitted essays addressing the contest topic, and they selected this paper as the winner.

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.


About the Author:
Joel Nolette is a litigation associate at a law firm in Boston, Massachusetts. He received his J.D. from Georgetown University Law Center in 2017.

I am grateful to William R. Levi, under whose guidance on the Senate Judiciary Committee I had the opportunity to mull over administrative law reform. I am also grateful to Nathan Kaczmarek and the Federalist Society’s Article I Initiative for providing the impetus to formulate this idea into an essay. Additionally, my thanks go to Katie McClendon, whose keen editorial sensibilities made this Article clearer and more precise before publication. Of course, the views presented in this Article are solely attributable to me, and all errors and mistakes are my fault alone.

“All legislative Powers herein granted shall be vested in a Congress of the United States . . . .”
—U.S. Constitution

“[T]o make government work for us, we need a Congress whose members are responsible for the consequences their decisions impose on us.”
—David Schoenbrod

“I will be honest with you. I do want the credit without any of the blame.”
—Michael Scott

Shortly after the Supreme Court struck down the legislative veto in Immigration and Naturalization Service v. Chadha, Representative Trent Lott observed:

We are often dumbfounded at how the law[s] we have enacted have been translated into regulations which have the force of law. In some cases the regulations may be within the scope of our enactments but we frankly did not anticipate what costs and burdens might be necessary to implement our intent. But, in other cases the regulations bear little resemblance to what we thought was our intent in passing those laws. In any case, we still bear the ultimate responsibility for these decisions, no matter how much we might try to pass the buck and scapegoat the bureaucrats. . . . [I]f we bear the ultimate responsibility for these decisions, we should exercise that responsibility before significant damage is done.

Congress did not heed this advice. The D.C. Circuit spoke of a “familiar” phenomenon in 2000: Congress passes vague or broadly worded statutes, and agencies implement them via regulations, interpretive guidance, and the like, essentially making law in Congress’ stead. With this system, Congress has figured out a way to “get the most credit for the least blame”—pass an underbaked law with an anodyne name, march out in front of the public to take credit for doing something about the issue in question, then blame the agency when the inevitable consequences of implementation—trade-offs, hard choices, winners and losers—foment political outrage. Worse yet, as Representative
Lott mentioned, the agency will begin implementing the law in ways Congress never intended or authorized. When this happens, congressional committees routinely will hale agency heads before them to chastise them for the policy choices they have made or for the mission creep occurring under their supervision. However, despite the fact that it crafted the law in question and has the power to amend it, Congress suddenly becomes unwilling or unable to do anything about it. Additionally, even where majorities in both chambers of Congress might want to change the law, procedural hurdles in the legislative process and the President's veto power effectively prevent Congress from reining in administrative agencies that have overstepped their bounds or made policy choices that Congress did not authorize.

This is not the way things are supposed to be. The government structure established by the Constitution is meant to protect liberty, securing freedom “both distinct from and every bit as important as those freedoms guaranteed by the Bill of Rights.”

The first structural constraint established by the Constitution is the vesting of all legislative powers in one Congress—powers that can only be exercised (for the most part) via specified procedures. The practice described above, however, circumvents this constraint: agencies promulgate rules “properly . . . regarded as legislative in . . . character and effect” without any formal process, let alone the constitutionally prescribed one of bicameralism and presentment.

Besides being in tension with the Framers’ constitutional design, these practices and procedures are at odds with our democratic traditions:

[I]n America . . . . [b]efore you can impose your views on the polity, you have to convince your fellow citizens that you’re right. That’s what democracy is all about. So it makes good sense to require the president to gain the support of Congress even when his vision is morally compelling. He should not be allowed to lead the nation on a great leap forward through executive decree.

Legislation-by-delegation has led to extensive executive lawmaking, frequently to the delight of allies in Congress. Yet what one executive does unilaterally, another can undo unilaterally. The pen-and-phone strategy of public policy can only guarantee ephemeral change. Such instability is corrosive to the rule of law.

The Supreme Court has largely absolved itself of responsibility for being Congress’ keeper in this regard, which is understandable to some extent: if Congress refuses to exercise its constitutional prerogatives, how is the Court going to force it to do so? It has limited options, and none seem promising. For instance, the Court could stop deferring to agency interpretations of ambiguous statutes and regulations, as some Justices have suggested, instead of presuming that ambiguity equals delegation and policy discretion. But this might not accomplish very much. After it was stripped of the unicameral legislative veto in Chadha, instead of reining from delegating power to agencies to make policies pursuant to broadly-worded statutes, Congress largely “abdicat[e]d its law-making function to the Executive Branch and independent agencies.” In similar fashion, Congress could react to the abandonment of Chevron and Seminole Rock deference simply by making broad but explicit delegations of lawmaking power and policy discretion to agencies, up to the outer limits of the nondelegation doctrine. The Court could put teeth back into the nondelegation doctrine, but the Justices have shown little interest in doing so. To boot, reinvigoration of the nondelegation doctrine might be difficult to square with the Constitution of the framers’ design. Maybe the time has come to face the behemoth.


See Chadha, 462 U.S. at 968 (White, J., dissenting).

Additionally, abrogating Seminole Rock alone, without accompanying reforms, could prove counterproductive and exacerbate the problems of fair notice and due process that have led to calls for reform. See Aaron L. Nielson, Beyond Seminole Rock, 105 Geo. L.J. 943, 947 (2017) (observing that “overruling Seminole Rock . . . may harm the very people” opponents of the doctrine “hope to help” because agencies have discretion under SEC v. Chenery, 332 U.S. 194 (1947) (Chenery II), to decide whether to promulgate rules by rulemaking, with various ex ante procedural safeguards, or by adjudication, with far fewer such protections); see also David L. Franklin, Legislative Rules, Nonlegislative Rules, and the Peril of the Short Cut, 120 Yale L.J. 276, 306–07 (2010) (“If policymaking by rule becomes sufficiently costly . . . then agencies will shift to purely adjudicatory mechanisms,” which “offers less notice and less opportunity for widespread participation.”).


Shortly after this was written, the Court granted certiorari to determine whether a federal statute’s “delegation of authority to the Attorney General to issue regulations . . . violates the nondelegation doctrine.” Gundy v. United States, 86 U.S.L.W. 5441, 2018 U.S. LEXIS 17
doctrine in its classical form might not make much of a difference in practice: the line between permissible and impermissible delegations is in the eye of the beholder.\textsuperscript{19} For a more effective and enduring solution to the unsustainable status quo, we need to find a different path up the mountain.\textsuperscript{20}

The thesis of this article is that such a solution is already at hand—almost hiding in plain sight. Borrowing from the rule of lenity and the contra proferentem doctrine, the Court should adopt a canon of strict construction for interpreting and applying federal statutes that delegate authority to agencies and regulations adopted pursuant to those statutes. Under this canon, an agency action\textsuperscript{21} will lack force or effect unless it is unambiguously authorized by statute or actually implements the unambiguous terms of a previously promulgated regulation. To determine whether a statute or regulation is unambiguous—and thus whether it actually authorizes the action in question instead of just leaving the matter unclear—\textsuperscript{22} the Court would use traditional tools of statutory interpretation.\textsuperscript{23} If the Court thereby concludes the statute or regulation is ambiguous, then it will construe the provision against the agency. This would replace the Court’s current practice of presuming an agency’s interpretation to be valid unless it is patently wrong.\textsuperscript{24} In other words, the Court would flip the presumptions of \textit{Chevron} and \textit{Seminole Rock}\textsuperscript{—if a law or rule is ambiguous, then the “tie” goes to the regulated rather than the regulator.\textsuperscript{25}}

This canon—dubbed here an “administrative rule of lenity”\textsuperscript{26}—could change the current dynamic that encourages

\begin{itemize}
  \item At least twice recently, something approaching this idea has been mentioned, albeit in passing. Professors Cass Sunstein and Adrian Vermeule, in an earlier draft of a recent article of theirs, identified as an alternative to \textit{Auer} deference the possibility “that in the face of ambiguity, the private sector is allowed to do what it wants.” Cass R. Sunstein & Adrian Vermeule, \textit{The Unbearable Rightness of Auer}, ___ U. Chi. L. Rev. ___, at 1-3 (Working Draft May 15, 2016), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2716737 [https://perma.cc/7XHL-F2TX]. They called this idea a “creative answer” with “some appeal” in some cases, albeit one that to their knowledge “has no defenders.” Id. Also recently, during a panel discussion at an administrative law conference, Professor Chris Walker identified as a “penalty for poor drafting” the proposal that courts should just “say no” whenever “it is clear that an agency is trying to expand its authority or is interpreting a statute in an aggressive way.” \textit{Separation of Powers: Congress, Agencies, and the Courts}, Rethinking Judicial Defection: History, Structure, and Accountability, Center for the Study of the Administrative State (June 2, 2016), available at https://vimeo.com/169757569 [https://perma.cc/FT4T-SPYT]; ORA Administrator Neomi Rao, also on the panel, remarked that “maybe there’s something to that” idea.

  \item Additionally, something akin to this idea may have at one time been advanced and rejected. See Landis, supra note 23, at 49 n.2 (quoting from an SEC opinion in which the Commission concluded that “[i]n the context of our powers under our fundamental Act with undue strictness at this stage in our growth would be to sacrifice upon the altar of a by-gone legal formalism our ability to perform adequately our allotted task”).

  \item Finally, at least two states have codified provisions that could lead to this result in some cases. Colorado law provides that “[n]o rule shall be issued except within the power delegated to the agency and as authorized by law,” and “[a] rule shall not be deemed to be within the statutory authority of any agency merely because such rule is not contrary to the specific provisions of a statute.” Colo. Rev. Stat. § 24-4-103(8)(a). Florida law states that “[a] grant of rulemaking authority is necessary but not sufficient to allow an agency to adopt a rule; a specific law to be implemented is also required.” Fla. Stat. § 120.536(1). That provision goes on, “[a]n agency may adopt only rules that implement or interpret the specific powers and duties granted by the enabling statute,” and “[n]o agency shall have the authority to adopt a rule only because it is reasonably related to the purpose of the enabling legislation and is not arbitrary and capricious or is within the agency’s class of powers and duties, nor shall an agency have the authority to implement statutory provisions setting forth general legislative intent or policy.” Id. Most relevant here, the statute then reads, “Statutory language granting rulemaking authority or generally describing the powers and functions of an agency shall be construed to extend no further than implementing or interpreting the specific powers and duties conferred by the enabling statute.” Id.

  \item Though titled an “administrative rule of lenity” here, this idea can just as readily be conceptualized as a “separation of powers clear statement” rule, cf Gregory v. Ashcroft, 501 U.S. 452, 460-61 (1991) (cleaned up) (discussing the federalism “plain statement rule” where, “if Congress intends to alter the usual constitutional balance between the States and the Federal Government, it must make its intention to do so unmistakably clear in the language of the statute”), or a “most questions” (as opposed to just “major questions”) doctrine, see King v. Burwell, 135 S. Ct. 2480, 2489 (2015) (internal quotation marks omitted) (discussing the rule in cases involving a “question of deep economic and political significance,” where the Court presumes that “had Congress wished to assign that question to an agency, it surely would have done so expressly”), as this proposal stems both from due process considerations.
\end{itemize}
government, and it briefly shows, through two contemporary examples, how this rule might operate in practice.

I. A Government of Laws and Not of Men

Foreshadowing the structure of the Federal Constitution that would follow along shortly after, the Massachusetts Constitution declares:

In the government of this commonwealth, the legislative department shall never exercise the executive and judicial powers, or either of them: the executive shall never exercise the legislative and judicial powers, or either of them: the judicial shall never exercise the legislative and executive powers, or either of them: to the end it may be a government of laws and not of men.

This design was not arbitrary. The Framers viewed separation of powers “as the absolutely central guarantee of a just Government.” At the heart of our system of government lies this core structural demarcation: the legislature makes the laws, the executive enforces the laws, and the judiciary interprets the laws. Given this separation of powers, “the legislative power of Congress cannot be delegated”—just as no one branch can assume the powers vested in another.

The Framers were not content “to trust . . . parchment barriers against the encroaching spirit of power” that might imperil this design. There needed to be “some practical security” to secure each branch “against the invasion of the others,” lest “a tyrannical concentration of all the powers of government in the same hands” result. Therefore, the Framers tried to design “the interior structure of the government as that its several constituent parts may, by their mutual relations, be the means of keeping each other in their proper places.” To this end, “the great security against a gradual concentration of the several powers in the same department, consists in giving to those who administer each department the necessary constitutional means and personal motives to resist encroachments of the others.”

Through this design, “the private interest of every individual”

---

**Footnotes:**

27 Some may, at this point, cry foul and marshal forth a parade of horribles about how this would gridlock government to a halt or require the legislature to speak at an impossibly precise level of specificity in statutes. See Mistretta, 488 U.S. at 372 (“It is often said in our increasingly complex society, replete with ever changing and more technical problems, Congress simply cannot do its job absent an ability to delegate power under broad general directives.”); Landis, supra note 23, at 2 (“The insistence upon the compartmentalization of power along triadic lines gave way in the nineteenth century to the exigencies of governance.”). Recent empirical research shows that these fears are unfounded. As Professor Walker has shown, “[f]ederal agencies help draft statutes,” both “in the foreground of the legislative process” and “in the shadows,” Christopher J. Walker, Legislativing in the Shadows, 165 U. PENN. L. REV. 1377, 1378–79 (2017). This being the case, it is not too much to ask for clarity, even where the subject matter might be technical. When elected legislators might not know what to say exactly, all they have to do is ask the technocrats, as they already do.

28 *Chenery II*, 332 U.S. at 217 (Jackson, J., dissenting).

29 Mass. Const. art. XXX.


32 *Cf. United States v. Welden*, 377 U.S. 95, 117 (1964) (Douglas, J., dissenting) (“Congress is not a law enforcement agency; that power is entrusted to the Executive. Congress is not a trial agency; that power is entrusted to the Judiciary.”).

33 *The Federalist No. 48*, at 305 (James Madison) (Clinton Rossiter ed., 1961).

34 Id. at 305, 310.


36 Id. at 318–19 (emphasis added).

37 Id. at 319.
would act as "a sentinel over the public rights." The mutually conflicting self-interest of those in power would prevent undue commingling of government functions, which, in turn, would safeguard freedom. This "separation of governmental powers ... is essential to the preservation of liberty." This design "was adopted ... not to promote efficiency but to preclude the exercise of arbitrary power" and "to save the people from autocracy." As the Chadha Court said:

The choices ... made in the Constitutional Convention impose burdens on governmental processes that often seem clumsy, inefficient, even unworkable, but those hard choices were consciously made by men who had lived under a form of government that permitted arbitrary governmental acts to go unchecked. There is no support in the Constitution or decisions of this Court for the proposition that the cumbersoness and delays often encountered in complying with explicit constitutional standards may be avoided, either by the Congress or by the President. Per this design, law, to be enforceable, must generally go through the gauntlet of the legislative process. This system cannot be short-circuited in the name of necessity, utility, practicality, or anything else—at least in theory.

II. The Rise and Rise of the Administrative State

The Founders viewed the legislature as the most dangerous branch. To Madison, "The legislative department is everywhere extending the sphere of its activity, and drawing all power into its impetuous vortex." For this reason, the Framers divided the federal legislature into two chambers "as little connected with each other as the nature of their common functions and their common dependence on the society [would] admit." As this partition shows, "the Framers were determined that the legislative power should be difficult to employ." They sought to preserve liberty by requiring the legislature to follow a rigorous process that would facilitate deliberation and compromise, and by calling for elections frequently enough to enable voters to hold representatives accountable. Chief among these procedural constraints is the requirement that a "matter which is properly to be regarded as legislative in its character and effect" go through the process of "bicameralism and presentment" before becoming law. This serves to "assure[] that the legislative power would be exercised only after opportunity for full study and debate in separate settings." This process also "promotes caution" and "rais[es] the decision costs of passing any law." In short, before law can be made, there must be sufficient consensus as to what the law should be, as well as cooperation between various factions and government actors with differing interests.

Things have changed. The three branches of government have effectively reconfigured our constitutional structure "in ways so fundamental as to suggest that something akin to a constitutional amendment ha[s] taken place." Congress now makes more lawmakers than laws by empowering bureaucrats in manifold administrative agencies to "fill in ... gaps" left by ambiguities in statutes. The executive has accumulated significant power through this process, and the Supreme Court has by and large endorsed this "more pragmatic, flexible approach" in light of "the contemporary realities of our political system." This reconfiguration has essentially made the executive branch the chief lawmaking body in the United States. Bureaucrats promulgate "legislative rules" which "have the 'force and effect of law'"—that is, they make law, notwithstanding

38 Id.
39 Mistretta, 488 U.S. at 380.
41 462 U.S. at 959.
42 Cf. The Federalist No. 78, at 464 (Alexander Hamilton) (Clinton Rossiter ed., 1961) (describing the judiciary as "the least dangerous to the political rights of the Constitution").
43 The Federalist No. 48, at 306 (James Madison) (Clinton Rossiter ed., 1961); see also The Federalist No. 51, at 319 (James Madison (Clinton Rossiter ed., 1961) (“In republican government, the legislative authority necessarily predominates.”).
45 Consumer Energy Council of Am. v. FERC, 673 F.2d 425, 464 (D.C. Cir. 1982); see also John F. Manning, Lawmaking Made Easy, 10 Green Bag 2d 191, 202, 204 (2007) (noting that the Constitution "manifestly places value upon cumbersoness, high transaction costs, and even ... gridlock").
46 Chadha, 462 U.S. at 952.
47 Id. at 951. See also Bowsher v. Synar, 478 U.S. 714, 722 (1986) (observing that bicameralism and presentment "assure full, vigorous, and open debate on the great issues affecting the people").
50 There are currently 441 listed federal agencies at the Federal Register website. See Federal Register, https://www.federalregister.gov/agencies [https://perma.cc/L5SH-MA58] (last visited Apr. 8, 2018).
52 Cf. Sunstein, supra note 49, at 447 (“[T]he repudiation of the system of separation and of checks and balances was a central feature of the New Deal reformation. By creating a new set of autonomous administrative actors, the New Deal critics sought to bypass the common law courts and, occasionally, the legislative process . . . .
54 Mortg. Bankers Ass’n, 135 S. Ct. at 1203 (quoting Chrysler Corp. v. Brown, 441 U.S. 281, 302–03 (1979)).
euphemistic parlance\textsuperscript{55} and despite assurances to the contrary.\textsuperscript{56} These legislative rules are then fleshed out in so-called “interpretive rules,” which are “issued by an agency to advise the public of the agency’s construction of the statutes and rules which it administers.”\textsuperscript{57} These interpretive rules formally “do not have the force and effect of law,”\textsuperscript{58} but, in reality, as long as they are not “plainly erroneous or inconsistent with the regulation,”\textsuperscript{59} interpretive rules “do have the force of law.”\textsuperscript{60} “Law is made” with scant procedural constraint.\textsuperscript{61} Instead of bicameralism and presentment, we have “agency lawmaking on the cheap.”\textsuperscript{62}

Thus obtains what Gary Lawson called “the rise and rise of the administrative state.”\textsuperscript{63} Back in 1952, Justice Jackson could write:

The rise of administrative bodies probably has been the most significant legal trend of the last century and perhaps more values today are affected by their decisions than by those of all the courts, review of administrative decisions apart. They also have begun to have important consequences on personal rights. They have become a veritable fourth branch of the Government, which has deranged our three-branch legal theories much as the concept of a fourth dimension unsettles our three-dimensional thinking.\textsuperscript{64}

These developments have continued apace, and the administrative state now “wields vast power and touches almost every aspect of daily life.”\textsuperscript{65} One can hardly go a day without being affected, directly or indirectly, by at least one of the alphabet soup of federal agencies: “agency rulemaking powers are the rule rather than, as they once were, the exception.”\textsuperscript{66}

Of course, no one forced the executive to claim the power the legislature chose to give (or at least left for the taking). But neither the President nor the bureaucrats alone can be faulted. Congress has willingly ceded its constitutional prerogatives, and the Supreme Court has placed its imprimatur on this state of affairs.

First, elected officials in Congress discovered it was easier to draft ambiguous laws and let the other branches sort it out than to risk political fire for making difficult legislative choices themselves. “Voters routinely punish lawmakers who . . . challenge them to face unpleasant truths,”\textsuperscript{67} so, assuming legislators want to be relected, it makes rational sense for them to “avoid[] hard choices by using general language and delegating to agencies the job of promulgating and implementing regulations.”\textsuperscript{68} So Congress passes aspirational statutes,\textsuperscript{69} delegating to unelected bureaucrats the difficult task of fleshing out the legislation outside of the legislative process. For example, the Clean Water Act, enacted in 1972, prohibits “the discharge of any pollutant by any person.”\textsuperscript{70} “Discharge” is defined as “any addition of any pollutant to navigable waters from any point source.”\textsuperscript{71} “[N]avigable waters” is then defined as “the waters of the United States.”\textsuperscript{72} What are “the waters of the United States”? Congress did not define the term, nor has it amended the law to provide a definition, despite the fact that the reach of the Clean Water Act remains a hotly contested issue to this day.\textsuperscript{73}

Second, the judiciary has aided and abetted the legislature’s decision to leave statutes ambiguous, allowing agencies to run

\textsuperscript{55} See Dept’ of Transp. v. Ass’n of Am. R.R., 135 S. Ct. 1225, 1246 (2015) (Thomas, J., concurring in the judgment) (“Implicitly recognizing that the power to fashion legally binding rules is legislative, we have nevertheless classified rulemaking as executive (or judicial) power when the authorizing statute sets out ‘an intelligible principle’ to guide the rulemaker’s discretion.”); City of Arlington v. FCC, 569 U.S. 290, 312, 315 (2013) (Roberts, C.J., dissenting) (“[M]odern administrative agencies . . . exercise legislative power . . . .”); cf. United States v. Seluk, 691 F. Supp. 525, 527–28 (D. Mass. 1988) (“[A]dvocates of contrasting views predictably use different terminology with contrasting tendencies as hidden persuaders. Characterizing the power . . . as a ‘legislative’ power encourages one to conclude that this power belongs only in the legislative branch. Similarly, on the other side, characterizing this power as a ‘rulemaking’ power encourages one to conclude that it is an inherent power of every court, administrative agency, and commission.”).

\textsuperscript{56} See Ernst & Ernst v. Hochfelder, 425 U.S. 185, 213 (1976) (“The rulemaking power granted to an administrative agency charged with the administration of a federal statute is not the power to make law.”).

\textsuperscript{57} Shalala v. Guernsey Memorial Hospital, 514 U.S. 87, 99 (1995).

\textsuperscript{58} Id.

\textsuperscript{59} Seminole Rock, 325 U.S. at 414.

\textsuperscript{60} Morg. Bankers Ass’n, 135 S. Ct. at 1212 (Scalia, J., concurring in the judgment).

\textsuperscript{61} Appalachian Power Co. v. EPA, 208 F.3d 1015, 1020 (D.C. Cir. 2000).

\textsuperscript{62} Manning, supra note 48, at 240.


\textsuperscript{64} FTC v. Ruberoid Co., 343 U.S. 470, 487 (1952) (Jackson, J., dissenting) (citation omitted).


\textsuperscript{66} Antonin Scalia, Judicial Deference to Administrative Interpretations of Law, 1989 Duke L.J. 511, 516.


\textsuperscript{68} Robert J. Pushaw, Jr., Talking Textualism, Practicing Pragmatism: Rethinking the Supreme Court’s Approach to Statutory Interpretation, 51 Ga. L. Rev. 121, 151 (2016).


\textsuperscript{70} 33 U.S.C. § 1311(a).

\textsuperscript{71} Id. § 1362(12).

\textsuperscript{72} Id. § 1362(7).

to “the outer limits” of federal power. 74 As long as Congress “lay[s] down by legislative act an intelligible principle to which the person or body authorized . . . is directed to conform, such legislative action is not a forbidden delegation of legislative power.”75 The Court has “almost never” challenged “Congress regarding the permissible degree of policy judgment that can be left to those executing or applying the law.”76 Instead, it has upheld the most generic delegations imaginable, including one instructing an agency to regulate whenever “public convenience, interest, or necessity will be served thereby.”77 Justice Thomas has observed that:

[T]he level of specificity [the Court] has required has been very minimal indeed. Under the guise of the intelligible-principle test, the Court has allowed the Executive to go beyond the safe realm of factual investigation to make political judgments about what is “unfair” or “unnecessary.” It has permitted the Executive to make trade-offs between competing policy goals. It has even permitted the Executive to decide which policy goals it wants to pursue. And it has given sanction to the Executive to craft significant rules of private conduct.78

This legislative abdication upends the expectations of the Framers, who took steps to “fortify” the other branches out of concern that the legislature would dominate the federal government.79 The risk of encroachment now stems from another place: “the citizen confronting thousands of pages of regulations—posed by the growing power of the administrative state cannot be given sanction to the Executive to craft significant rules of private conduct.”78

The level of specificity [the Court] has required has been very minimal indeed. Under the guise of the intelligible-principle test, the Court has allowed the Executive to go beyond the safe realm of factual investigation to make political judgments about what is “unfair” or “unnecessary.” It has permitted the Executive to make trade-offs between competing policy goals. It has even permitted the Executive to decide which policy goals it wants to pursue. And it has given sanction to the Executive to craft significant rules of private conduct.78

This legislative abdication upends the expectations of the Framers, who took steps to “fortify” the other branches out of concern that the legislature would dominate the federal government.79 The risk of encroachment now stems from another place: “the citizen confronting thousands of pages of regulations—posed by the growing power of the administrative state cannot be

81 Id.

82 See Nielson, supra note 17, at 998 (“The line between ‘change’ and ‘clarification’ is . . . a question of degree more than kind.”).

83 See Texas v. United States, 787 F.3d 733, 746 (5th Cir. 2015) (involving a challenge to President Obama’s “deferred action” programs for individuals in the country unlawfully, which the administration contended were “exempt from . . . notice-and-comment”).


85 5 U.S.C. § 551 et seq.

86 Even those sympathetic to the policy agenda recognize the problems with this process. See, e.g., Sam Williamson, Note & Comment, G.G. ex rel. Grimm v. Gloucester Cty. Sch. Bd.: Broadening Title IX’s Protections for Transgender Students, 76 Md. L. Rev. 1102, 1103 (2017) (footnotes omitted) (“While the Fourth Circuit’s decision led to the appropriate conclusion—that the Gloucester County School Board must give Gavin access to the boys’ restroom—the Fourth Circuit should have conducted its own statutory interpretation . . . . The judiciary should not continue to grant deference to administrative agencies on questions concerning the rights of discrete and insular minorities when the agencies’ authority to regulate was merely implicitly delegated.”).


At this point, some “critics of the critics,” as it were—that is, those who think the critics of the administrative state overstate their case—might respond that fears of agency gamesmanship and wholesale end-runs around either legislative or administrative processes are overblown, either because courts already tend to do a good job of checking agency mission creep or because agencies do not engage in such mischief all that often. See, e.g., Robin Alexander Smith, Perez v. Mortgage Bankers Association and the Future of Seminole Rock, 40 Harv. Envtl. L. Rev. 173, 181 (2016) (internal quotation marks omitted) (noting that courts of appeals have not deferred to agency interpretation in cases where “it provides the basis for an enforcement action, has an immediate or direct effect on regulated parties, or is otherwise viewed as controlling in the field”); Connor N.
about this last point only when their political opponents occupy the White House. But—we—should always care.

III. TOWARDS AN ADMINISTRATIVE RULE OF LENILITY

The ultimate solution to the problems created by congressional self-abnegation is congressional self-assertion. In an ideal world, an enlightened citizenry would elect representatives willing to tell them hard truths rather than comforting lies, and these representatives would do their job in the legislature by dealing with difficult public policy issues rather than shunting them off for resolution in the opaque administrative state. But "experience" here has taught us "the necessity of auxiliary precautions" in addition to those structural checks the Framers built into the Constitution. But what can possibly work to save Congress from itself, especially when it apparently would prefer to be saved by the other branches?

To start, the executive could order agencies under its control to stop making law by informal guidance and similar means. This, however, is not a lasting fix, as a new administration could easily undo it. To get closer to effecting real change, the Court could adopt a rule of construction that would shift legislative incentives and require Congress to speak clearly when making law—the "administrative rule of lenity."

Under this rule, the Court would refuse to give effect to any agency action that Congress did not clearly authorize in a duly-enacted statute. By essentially flipping the default rules embodied in Chevron and Seminole Rock, this rule could alter the dynamic among Congress, agencies, and the electorate that has led to the status quo. It would strip Congress of the political benefits obtained from delegating away legislative power in broad, vague terms, because Congress could no longer get away with failing to make policy choices itself but then faulting (or praising, as the case may be) agencies for their policy choices. If federal policy binding on the citizenry is to be made, Congress is going to have to make it. This rule would in turn encourage agencies or the executive to seek congressional approval by way of new legislation any time there is a risk that contemplated agency action might exceed the scope of existing law. Perhaps most important of all, this would ensure that individuals are subject only to duly-enacted laws (and fixed judicial constructions thereof), not agency decree.

While there may be times when Congress can properly make a "general provision" and leave it "to those who are to act under such general provisions to fill up the details," what qualifies as a

---

89 See James L. Buckley, Saving Congress from Itself: Emancipating the States & Empowering Their People (2014); see also Gerald G. Ashdown, Marshall, Marbury, and Mr. Byrd: America Unchecked and Imbalanced, 108 W. Va. L. Rev. 691, 703 (2006) ("Congress will have to protect itself. Byrd thinks they are losing the battle, sometimes with self-inflicted wounds.").

90 See House of Representatives v. Burwell, 130 F. Supp. 3d 53, 57 (D.D.C. 2015) ("[T]he House of Representatives complains that . . . the Secretary of Health and Human Services . . . the Secretary of the Treasury, and their respective departments . . . have spent billions of unappropriated dollars to support the Patient Protection and Affordable Care Act."). To be fair, things have become so distorted that even when Congress tries to defund executive action, the executive branch has recourse to workarounds. See House of Representatives v. Burwell, No. 14-1967 (RMC), 2016 WL 2750934, at *6, *18 (May 12, 2016) (discussing how the Treasury Department made reimbursement payments "to issuers of qualified health plans" under the Affordable Care Act despite the fact that Congress did not "appropriate[] money for . . . reimbursements").

91 See Attorney General Jeff Sessions Ends the Department’s Practice of Regulation by Guidance, Dep’t of Justice (Nov. 17, 2017), available at https://www.justice.gov/opa/pr/attorney-general-jeff-sessions-ends-department-practice-regulation-guidance [https://perma.cc/L1J4-UHDA] ("Today, in an action to further uphold the rule of law in the executive branch, Attorney General Jeff Sessions issued a memo prohibiting the Department of Justice from issuing guidance documents that have the effect of adopting new regulatory requirements or amending the law. The memo prevents the Department of Justice from evading required rulemaking processes by using guidance memos to create de facto regulations.").

detail is itself no detail."94 Determining what is a matter of policy for Congress and what are details that can be left to agencies is a difficult line-drawing exercise, and the lines drawn are likely to shift with the political winds or the composition of the Court.95 A clear background rule like this—functioning like Chevron, although cutting in the opposite direction of Chevron—would promote clarity and specificity.96

The Court need not look far to justify such an approach. It already has useful analogues at hand to support its adoption of such a rule, and those analogues have similar supporting rationales. The criminal law rule of lenity and the contract law doctrine of contra proferentem illustrate why an administrative rule of lenity is both prudent and practical.

A. The Rule of Lenity

In criminal law, the rule of lenity requires a court to construe an ambiguous criminal statute in favor of the defendant.97 This rule puts the due process principle of fair notice into practice, protecting people from liability for crimes they could not have known were crimes.98 Lenity also creates a dynamic in which the executive is encouraged to “induce Congress to speak more clearly” so that enforcement will not be hindered due to imprecise statutory language.99 Additionally, it forces Congress to “lay bare the full extent of the conduct they intend to” proscribe or penalize, “exposing themselves to whatever resistance or ridicule their choices entail” and ensuring “that the public had maximum opportunity to influence . . . the enactment of the prohibition . . . .”100

This rule has a venerable pedigree: courts have historically “refused to apply” vague laws “under the rule that penal statutes should be construed strictly.”101 Yet “no one contends that the rule of lenity should apply in the civil context” where property rights, but not personal liberty, are at stake, as is frequently the case in administrative disputes.102 The usual explanation for the differential application of the rule of lenity is that the Court has “expressed greater tolerance of enactments with civil rather than criminal penalties because the consequences of imprecision are qualitatively less severe.”103 This rationalization falls apart upon examination.

In the text of the Constitution, due process applies equally to deprivations of life, liberty, and property, without differentiation. Nevertheless, courts have reasoned that the rule of lenity need not apply to “an indefinite civil statute” like it does to a criminal one because it is “a more serious matter to deprive a man of his liberty on a prosecution based upon a vague and indefinite statute than to deprive him of a property right alone.”104 But the Court has

94 United States v. Nichols, 784 F.3d 666, 671 (10th Cir. 2015) (Gorsuch, J., dissenting from the denial of rehearing en banc).
95 See supra note 19 and accompanying text.
96 Cf. Scalia, supra note 66, at 517 (discussing the merits of “a background rule of law against which Congress can legislate”).
97 See Bell v. United States, 349 U.S. 81, 83 (1955) ("When Congress leaves to the Judiciary the task of imputing to Congress an undeclared will, the ambiguity should be resolved in favor of lenity.").
98 United States v. Santos, 553 U.S. 507, 514 (2008) (plurality op.) (pointing out that the rule of lenity “indicates the fundamental principle that no citizen should be held accountable for a violation of a statute whose commands are uncertain”) (citation omitted); see Dunn v. United States, 442 U.S. 100, 112 (1979) (“[F]undamental principles of due process . . . mandate that no individual be forced to speculate, at peril of indictment, whether his conduct is prohibited.”); see also John M. Darley, Kevin M. Carlsmith, & Paul H. Robinson, The Ex Ante Function of the Criminal Law, 35 L. & Soc’y Rev. 165, 165 (2001) (“A legal code in a complex society is designed . . . . to announce beforehand the rules by which citizens must conduct themselves . . . .”).
99 Santos, 553 U.S. at 514 (citation omitted); see Liparota v. United States, 471 U.S. 419, 427 (1985) (noting that the rule of lenity “strikes the appropriate balance between the legislature, the prosecutor, and the court in defining criminal liability”); cf. Crandon v. United States, 494 U.S. 152, 160 (1990) (“Because construction of a criminal statute must be guided by the need for fair warning, it is rare that legislative history or statutory policies will support a construction of a statute broader than that clearly warranted by the text.”).
101 Johnson v. United States, 135 S. Ct. 2551, 2568 (2015) (Thomas, J., concurring in the judgment); see Crandon, 494 U.S. at 158 (referring to the rule of lenity as a “time-honored interpretive guideline”); see also Kevin C. Walsh, Partial Unconstitutionality, 85 N.Y.U. L. Rev. 738, 756 (2010) (“[W]hat we now call judicial review consisted of a refusal to give a statute effect as operative law in resolving a case.”).
102 Sash v. Zenk, 439 F.3d 61, 65 (2d. Cir. 2006) (Sotomayor, J.); see also Babbitt v. Sweet Home Chapter of Cmtyts. for a Great Or., 515 U.S. 687, 704 n.18 (1995) (“We have never suggested that the rule of lenity should provide the standard for reviewing facial challenges to administrative regulations whenever the governing statute authorizes criminal enforcement.”); Price, supra note 100, at 910 n.166 (“[T]he rule of lenity's method of resolving conflicts automatically in favor of the narrower view has no analogue in the interpretation of civil statutes . . . .”).

Some courts have questioned whether this is the case. See Hill v. Coggins, 867 F.3d 499, 514 (4th Cir. 2017) (“[I]t is unclear whether the rule of lenity applies in a civil dispute.”); Leslie Salt Co. v. United States, 55 F.3d 1388, 1398 (9th Cir. 1995) (“The rule of lenity has not been limited to criminal statutes, particularly when the civil sanctions in question are punitive in character.”); Whitfield v. United States, 99 A.3d 650, 656 n.14 (D.C. 2014) (citation omitted) (“The fact that [the law in question] is a civil traffic regulation, rather than an actual criminal statute, is of no moment. The rule of lenity is not so unduly restrictive in its application.”); United Pharmacal Co. of Mo., Inc. v. Mo. Bd of Pharm., 208 S.W.3d 907, 913 (Mo. 2006) (“The rule [of lenity] is applicable where violation of a civil statute has penal consequences.”); Rand v. Comm’r, 141 T.C. 376, 393 (T.C. 2013) (“[A]lthough often considered in the criminal context, the rule of lenity has been applied in the civil context and specifically with regard to civil tax penalties.”). However, these instances appear to be the exception rather than the rule.
104 General Const. Co. v. Connally, 3 F.2d 666, 667 (W.D. Okla. 1924); see also Spell v. McDaniel, 591 F. Supp. 1090, 1105 (E.D.N.C. 1984) (stating “that property deprivations are qualitatively different from deprivations of liberty” for purposes of the due process clause); cf. Edgar A. Levy Leasing Co. v. Siegal, 258 U.S. 242, 250 (1922) (affirming a civil law rule of lenity).
rebuffed this type of thinking. The Court, in fact, has called the "dichotomy between personal liberties and property rights" "a false one." Instead, the Court has recognized that "a fundamental interdependence exists between the personal right to liberty and the personal right in property," and "[n]either could have meaning without the other." Influential thinkers like John Locke even cast life and liberty as species of the right to property—"every Man has a Property in his own Person," which "no Body has any Right to but himself." This "false dichotomy" falls apart further when one considers situations "when a substantial deprivation of property will appear more egregious than a minimal intrusion on one's liberty." Having an across-the-board rule to require fair notice of "what the law intends to do if a certain line is passed" makes intuitive sense, whether the proscriptions and concomitant sanctions are criminal or civil.

Even if one accepts the liberty/property distinction that has been imported into the Due Process Clauses by many courts and commentators, a more expansive application of the rule of lenity finds further justification in light of the increasing "disappearance of any clearly definable line between civil and criminal law." With growing frequency, "much of the conduct that we would typically consider to be a violation of a regulation, subject to civil penalties in federal court or in an administrative tribunal, is criminalized in the same statute." Take again, for example, the Clean Water Act. In one section of the Act, the law imposes a criminal penalty of "no[] more than $25,000 per day of violation" for negligently breaking the law. In the very next subsection of the same code provision, the law imposes a civil penalty "not to exceed $25,000 per day for each violation" of the law. Should the rules governing statutory interpretation turn on "magic words or labels," when there is little distinction in reality or effect—in this case, simply the characterization of an otherwise identical penalty? One would think not: after all, a reviewing court should look to "substance and application" and consider a law’s "practical operation" rather than merely the "form of descriptive words which may be applied to it."

The Court has, in fact, already applied the rule of lenity outside a strictly criminal context. Because statutes must be construed consistently even when they have both criminal and civil applications, the Court has applied the rule of lenity when a law had such dual application, even if, in a given case, only civil penalties were threatened. Recently, Justice Ginsburg, joined by Justices Scalia, Sotomayor, and Kagan, called for application of the rule in a case where a law threatened "draconian civil liability." Language from earlier decisions supports the idea that lenity should be applied more broadly. Some state courts have overtly done so, applying a rule of strict construction when "statutes and rules are penal in nature," even if they threaten only "a civil penalty," including in cases involving "administrative law and procedures.

A wider sweep to lenity would be "consonant alike with ordinary notions of fair play and the settled rules of law, whether life, liberty, or property is threatened." It would encourage both Congress and agencies to speak clearly in the first instance, and it would urge agencies to seek change by influencing legislation.

---

108 See, e.g., Localio v. Ashcroft, 543 U.S. 1, 11 n.8 (2004) (“Because we must interpret the statute consistently, whether we encounter its application in a criminal or noncriminal context, the rule of lenity applies.”); see also United States v. Thompson/Center Arms Co., 504 U.S. 505, 518 (1992) (noting that the rule of lenity can be properly invoked in a civil case when dealing with a criminal statute); Crandon v. United States, 494 U.S. 152, 158 (1990) (“[B]ecause the governing standard is set forth in a criminal statute, it is appropriate to apply the rule of lenity in resolving any ambiguity in the ambit of the statute’s coverage.”).
112 John Locke, Second Treatise of Government § 27.
113 Frost v. City & Cty. of Honolulu, 584 F. Supp. 356, 361 (D. Hawai’i 1984); see infra notes 140–144 and accompanying text.
114 Id.
117 John C. Coffee, Jr., Does “Unlawful” Mean “Criminal”? Reflections on the Disappearing Tort/Crime Distinction in American Law, 71 B.U. L. REV. 193, 193 (1991); see also Mary M. Cheh, Constitutional Limits on Using Civil Remedies to Achieve Criminal Law Objectives: Understanding and Transcending the Criminal-Civil Law Distinction, 42 HASTINGS L.J. 1325, 1329 (1991) (“It is extremely difficult to draw principled lines to distinguish between criminal and civil cases.”).
118 Id. § 1319(c)(1).
119 In re Woodrow Wilson Constr. Co., 563 So.2d 385, 391 (La. Ct. App. 1990); see also People v. Mobil Oil Corp., 192 Cal. Rptr. 155, 164 (Cal. Ct. App. 1983) (“Because the statute is penal, we adopt the narrowest construction of its penalty clause to which it is reasonably susceptible in the light of its legislative purpose. This principle is not rendered inapplicable merely because an action arises out of an administrative proceeding rather than a criminal prosecution. Where the statute to be construed is a penal one, these principles apply even when the underlying action is civil in nature.”). But see Handymen Connection of Sacramento, Inc. v. Sands, 20 Cal. Rptr. 3d 727, 747 (Cal. Ct. App. 2004) (“Handymen urges us either to import the so-called ‘rule of lenity’ from criminal law into civil administrative law, or to treat occupational regulatory statutes as contracts of adhesion to which licensees are involuntary parties. To do either would be unwarranted.”).
rather than by “interpretation.” An administrative rule of lenity would also increase transparency and accountability: no longer could Congress delegate to an agency the task of balancing policy trade-offs pursuant to a mandate to, say, “regulate in the public interest,” without specifying what, exactly, constitutes “public interest” and what rules might promote that interest. Congress would have to make hard choices itself and speak clearly, or else risk non-enforcement of its vague enactments.

B. Contra Proferentem

In contract law, a court faced with an unresolvable ambiguity in an agreement construes the ambiguity against the party responsible for inserting it into the contract. This canon of construction, known as contra proferentem, “provides that [i]n choosing among the reasonable meanings of a promise or agreement . . . that meaning is generally preferred which operates against the party who supplies the words or from whom a writing otherwise proceeds.” This rule makes sense: “the drafter of a contract occupies an advantageous position regarding the language of the contract,” and “fairness requires as a matter of law that the bigger piece of the contract ‘pie’ not go to the slicer.”

Similarly, the legislature that writes a law or the agency that promulgates a regulation occupies an advantageous position relative to those who will be subject to those rules. If fairness requires the non-drafting party to benefit from an ambiguity in a contract between private parties, it makes sense that non-drafting parties in a coercive context like regulation should have ambiguities in laws or regulations construed in their favor as well.

As discussed above, this is especially true since agencies have means by which they can avoid any formal process altogether in effecting amendments to rules.

In the administrative state, however, the opposite of contra proferentem is the law. Under Seminole Rock, if “the meaning of the words used” in a regulation “is in doubt,” the promulgating agency’s construction “becomes of controlling weight unless it is plainly erroneous or inconsistent with the regulation.”

This, especially along with Chevron, has proved to be a very large thumb on the scale in favor of the government whenever a term in a statute or regulation is fairly susceptible of more than one interpretation. Agencies have exploited this advantageous playing field to adopt expansive interpretations of unclear statutes and regulations. The Court has recognized that “a deprivation of the right of fair warning can result not only from vague statutory language but also from an unforeseeable and retroactive judicial expansion of narrow and precise statutory language.”

But the Court has basically endorsed “unforeseeable . . . expansion of statutory or regulatory language by deferring to agencies’ interpretations of their own rules, giving those interpretations “the force of law” unless they are completely irrational.

Furthermore, like with contracts of adhesion that unsophisticated consumers neither read nor comprehend, it “is totally unrealistic to assume that more than a fraction of the persons and entities affected by a regulation . . . have knowledge of its promulgation or familiarity with or access to the Federal

123 See Price, supra note 100, at 888 (“The rule of lenity . . . compels lawmakers and enforcers to indicate explicitly what they are doing.”).

124 Cf. id. at 886 (noting how “legislatures” “routinely” delegate the responsibility of defining terms in statutes “to courts and executive officials,” but “[l]enience . . . blocks expansive readings and impedes delegated discretion by requiring courts to choose narrow interpretations automatically”).

125 Mesa Air Grp., Inc. v. Dept’ of Transp., 87 F.3d 498, 506 (D.C. Cir. 1996) (quoting Restatement (Second) of Contracts § 206 (1979)).


127 Cf., e.g., Sackett v. EPA, 566 U.S. 120, 132 (2012) (Alito, J., concurring) (“The reach of the Clean Water Act is notoriously unclear. . . . Any piece of land that is wet at least part of the year is in danger of being classified as wetlands covered by the Act, and according to the Federal Government, if property owners begin to construct a home on a lot that the Agency thinks possesses the requisite wetness, the property owners are at the Agency’s mercy. The EPA may issue a compliance order demanding that the owners cease construction, engage in expensive remedial measures, and abandon any use of the property. If the owners do not do the EPA’s bidding, they may be fined up to $75,000 per day ($37,500 for violating the Act and another $37,500 for violating the compliance order). . . . In a Nation that values due process, not to demand the owners construct a home on a lot that the Agency thinks possesses the requisite wetness, the property owners are at the Agency’s mercy. The EPA may issue a compliance order demanding that the owners cease construction, engage in expensive remedial measures, and abandon any use of the property. If the owners do not do the EPA’s bidding, they may be fined up to $75,000 per day ($37,500 for violating the Act and another $37,500 for violating the compliance order). . . . In a Nation that values due process, not to mention private property, such treatment is unthinkable. . . . Real relief requires Congress to do what it should have done in the first place: provide a reasonably clear rule regarding the reach of the Clean Water Act.”).

128 Mortg. Bankers Ass’n, 135 S. Ct. at 1211–12 (Scalia, J., concurring in the judgment) (observing that, thanks to “judge-made doctrines of deference,” agencies can use “interpretive rules,” which are exempt from ex ante notice-and-comment requirements, “not just to advise the public” regarding the meaning of existing law “but also to bind them” to that interpretation).

129 Seminole Rock, 325 U.S. at 413–14.

130 467 U.S. 837.

131 See William N. Eskridge, Jr. & Lauren E. Baer, The Continuum of Deference: Supreme Court Treatment of Agency Statutory Interpretations from Chevron to Hamdan, 96 Geo. L.J. 1083, 1104 (2008) (noting that “[t]he agency win rate for cases where the Court invoked Seminole Rock (or an analogous precedent) was an outstanding 90.9%”).

132 See, e.g., Sackett, 566 U.S. at 133 (Alito, J., concurring) (“Congress did not define what it meant by ‘the waters of the United States’; the phrase was not a term of art with a known meaning; and the words themselves are hopelessly indeterminate. Unsurprisingly, the EPA and the Army Corps of Engineers interpreted the phrase as an essentially limitless grant of authority.”).


134 Mortg. Bankers Ass’n, 135 S. Ct. at 1212 (Scalia, J., concurring in the judgment); see also Sackett, 566 U.S. at 132 (Alito, J., concurring) (observing that “the combination of the uncertain reach of federal law “and the draconian penalties imposed for the sort of violations alleged in this case . . . leaves most property owners with little practical alternative but to dance to” agencies’ “tune” regarding what the law says).

It will be of little avail to the people, that the laws are made by men of their own choice, if the laws be so voluminous that they cannot be read, or so incoherent that they cannot be understood; if they be repealed or revised before they are promulgated, or undergo such incessant changes that no man, who knows what the law is to-day, can guess what it will be to-morrow. Law is defined to be a rule of action; but how can that be a rule, which is little known, and less fixed?\textsuperscript{137}

Laws and regulations should “have sufficient content and definitiveness as to be a meaningful exercise”;\textsuperscript{138} if they do not, the reviewing court should construe ambiguity in favor of the non-drafting party, just as courts do already with contracts, especially given the imbalance in bargaining power between the regulators and the regulated. The administrative rule of lenity thus provides a bright-line background rule to avoid entirely “the potential due process problems posed by ‘penalizing a private party for violating a rule without first providing adequate notice of the substance of the rule.”\textsuperscript{139}

IV. Operationalizing the Administrative Rule of Lenity

How would applying this canon work in practice? Two examples—one involving a statute and one a regulation—illustrate how this might work to accomplish the intended result of encouraging Congress to draft laws with more precision and urging agencies to turn to Congress to make new law rather than doing so themselves through so-called interpretation.

A. Waters of the United States

Andy Johnson built a stock pond on his property in Wyoming by damming a small creek that ran through his front yard.\textsuperscript{140} Little did he know this act would ignite a “regulatory war”:\textsuperscript{141}

[O]fficials from the Environmental Protection Agency paid a visit to the pond and . . . told [Johnson] he was facing “a very serious matter.” In a January 2014 violation notice, the agency said Mr. Johnson had violated the Clean Water Act by digging out Six Mile Creek and dumping in tons of river rocks without getting necessary federal permits. The agency ordered him to take steps to restore the creek under the supervision of environmental officials, or face accumulating fines of as much as $37,500 a day.\textsuperscript{142}

Rejecting the EPA’s position, but unwilling either to destroy his stock pond or to incur ruinous fines with fingers crossed hoping eventually to be vindicated, Johnson sued in federal court. He argued, among other things, that the creek was “too far removed from navigable rivers to fall under the E.P.A.’s authority.”\textsuperscript{143} The case ultimately settled after a two-year struggle.\textsuperscript{144}

Consider how the case would have played out if the administrative rule of lenity were in effect. Because Congress left “waters of the United States” ambiguous, despite the fact that the phrase is notoriously unclear,\textsuperscript{145} Johnson could have responded confidently to the EPA’s compliance order by arguing that his creek could not be regulated as “waters of the United States,” particularly since it clearly was not “navigable.”\textsuperscript{146} If Congress wanted to give the EPA jurisdiction and enforcement authority over this creek, it would have to do so by defining “waters of the United States” so as to include such a body of water. Until that happened, or until a fixed and final judicial construction of the text were provided authorizing such a broad scope to the Act,\textsuperscript{147} Johnson could not be penalized under the Clean Water Act, thereby vindicating the principle that “no citizen should be held accountable for a violation of a statute whose commands are uncertain.”\textsuperscript{148}

With an administrative rule of lenity in place, this “regulatory war” could have been avoided before the first shot was fired. Congress is at fault for the ambiguous reach of the law, and until Congress fixes the problem or the Court fixes the law’s construction, no one should be penalized where the law’s reach is uncertain.


\textsuperscript{137} The Federalist No. 62, at 379 (James Madison) (Clinton Rossiter ed., 1961).

\textsuperscript{138} Paralyzed Veterans of Am. v. D.C. Arena L.P., 117 F.3d 579, 584 (D.C. Cir. 1997).

\textsuperscript{139} United States ex rel. Purcell v. MWI Corp., 807 F.3d 281, 287 (D.C. Cir. 2015) (quoting Satellite Broad. Co. v. FCC, 824 F.2d 1, 3 (D.C. Cir. 1987)).


\textsuperscript{141} Id.

\textsuperscript{142} Id.


\textsuperscript{144} Timothy Cama, EPA Settles with Wyoming Farmer over Man-Made Pond, This Hill (May 10, 2016), http://thishill.com/policy/energy-environment/279421-epa-settles-water-pollution-case-with-wyoming-farmer [https://perma.cc/8G7K-9BJ9] (“Under the settlement reached Monday in federal court, Johnson will not have to pay the fines or drain the pond.”).

\textsuperscript{145} See 33 U.S.C. § 1362(7); Santos, 566 U.S. at 133 (Alito, J., concurring) (“Congress did not define what it meant by ‘the waters of the United States’; the phrase was not a term of art with a known meaning; and the words themselves are hopelessly indeterminate.”).

\textsuperscript{146} 33 U.S.C. § 1362(7).

\textsuperscript{147} See Rapanos, 547 U.S. at 758 (Roberts, C.J., concurring) (“It is unfortunate that no opinion commands a majority of the Court on precisely how to read Congress’ limits on the reach of the Clean Water Act. Lower courts and regulated entities will now have to feel their way on a case-by-case basis.”).

\textsuperscript{148} Santos, 553 U.S. at 514 (plurality op.) (citations omitted).
B. On the Basis of Sex

Title IX of the Education Amendments of 1972 declares that “[n]o person . . . shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance,” and it empowers federal agencies “to effectuate the provisions” of the statute. The Department of Education has issued regulations doing so, one of which allows recipients of federal education funds to “provide separate toilet, locker room, and shower facilities on the basis of sex,” so long as the facilities are comparable. As the text of the law suggests, “Title IX was enacted in response to evidence of pervasive discrimination against women with respect to educational opportunities.” Regulations implementing Title IX must “effectuate [its] provisions”; an agency cannot, “under the guise of interpreting a regulation,” “create de facto a new regulation,” especially not one that sweeps beyond the scope of the underlying enabling act. Regulations, therefore, must comport with the goal of ending the unequal treatment of women in education.

On May 13, 2016, the Department of Justice and Department of Education jointly released a “Dear Colleague” letter announcing that Title IX’s prohibition on sex discrimination “encompasses discrimination based on a student’s gender identity.” This meant that, while “[a] school may provide separate facilities on the basis of sex” pursuant to agency regulations, schools across the country could “not require transgender students to use facilities inconsistent with their gender identity or to use individual-user facilities when other students are not required to do so,” at least not without jeopardizing federal funds.

This letter was hailed by some as a “historic step” for the Obama administration “in its aggressive defense of civil rights.” It was also “sure to stoke growing outrage across the country over what has become an explosive political issue.” Officials from the respective departments published a blog post in tandem with the letter, stating that “protecting transgender students’ right to be who they are does not harm other students,” and that this protection was “achievable through common-sense approaches that foster safety and a positive learning environment for all students.” Apparently, these officials also thought these goals were achievable without congressional action or even notice-and-comment rulemaking, despite the facts that “[w]hen Title IX was enacted in 1972, the term ‘sex’ was commonly understood to refer to the biological differences between males and females,” the early users of the term ‘gender identity’ recognized the distinction between ‘sex’ and ‘gender identity,’” and Title IX itself indicates a “binary definition” for “sex.”

The policy was rescinded before it could be litigated to the Supreme Court. Given the major, nationwide implications of this “guidance,” it is possible that the Supreme Court would not have deferred to the agencies in light of the fact that this is the sort of situation where, “had Congress wished to assign that question to an agency, it surely would have done so expressly.” But this result would have been even more certain under the administrative rule of leniency. That rule would have resulted in an automatic loss for the agencies in any challenge to their policy. For such a dramatic expansion of the scope of agency regulations to have any binding effect, Congress would have had to amend Title IX to explicitly authorize it. Until that happened, because neither Title IX nor its implementing regulations clearly indicate that “sex discrimination” includes “gender identity discrimination,” no agency could enforce such a position against an entity that disagrees.

“If an agency wants to exercise expansive regulatory authority over some major social or economic activity . . . an ambiguous grant of statutory authority is not enough. Congress must clearly authorize an agency to take such a major regulatory

---

150 Id. § 1682.
151 34 C.F.R. § 106.33.
152 McCormick v. Sch. Dist. of Mamaroneck, 370 F.3d 275, 286 (2d Cir. 2004); see also N. Haven Bd. of Educ. v. Bell, 456 U.S. 512, 523–24 (1982) (observing that Title IX was one of “several attempts . . . to enact legislation banning discrimination against women in the field of education” in the 1970s).
153 Christensen v. Harris Cty., 529 U.S. 576, 588 (2000). See also City of Arlington, 569 U.S. at 297 (“No matter how it is framed, the question a court faces when confronted with an agency’s interpretation of a statute it administers is always, simply, whether the agency has stayed within the bounds of its statutory authority.”); Massachusetts v. EPA, 549 U.S. 497, 533 (2007) (observing how agencies are confined “to exercise discretion within defined statutory limits”); Brown v. Gardner, 513 U.S. 115, 118 (1994) (“Ambiguity is a creature not of definitional possibilities but of statutory context.”).
155 Id.
158 Id.
161 King v. Burwell, 135 S. Ct. 2480, 2489 (2015). See also Stephen G. Breyer, Judicial Review of Questions of Law and Policy, 58 ADMIN. L. REV. 363, 370 (1986) ("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.")
In light of the foregoing principles, the same should be true whether an action is categorized as “major,” “minor,” or somewhere in between—especially since agency action falling anywhere between those two posts might still have “major” consequences for regulated parties. Before the executive may act to enforce a law, Congress must empower it to do so by actually enacting that law. The legislative process might pose significant obstacles to passing laws, and this might frustrate those seeking to advance a policy that they feel justice requires. But these obstacles are a feature, not a bug, of our Constitution.

V. Conclusion

The Constitution created a federal government with three separate branches “not to promote efficiency but to preclude the exercise of arbitrary power.” Congress, however, has largely abandoned its central, lawmaking role due to an incentive structure that encourages members to punt hard issues to the executive branch by delegating significant lawmaking power to administrative agencies. This has distorted the basic structure of the Constitution, weakened key constraints against arbitrary government, and eroded rule of law. The Supreme Court need not sit idly by while this happens: a canon of strict construction applied to federal administrative statutes and concomitant regulations could alter the playing field and force Congress to make laws rather than lawmakers by legislating in clear and specific terms. Adapting such a rule would not require a jurisprudential revolution—it would only require the Court to give fuller effect to extant rules whose rationale applies fittingly to this situation. Such a background principle could compel Congress to take responsibility anew to exercise the legislative power entrusted to it by “We the People.”

162 U.S. Telecom Ass’n v. FCC, 855 F.3d 381, 421 (D.C. Cir. 2017) (Kavanaugh, J., dissenting from denial of rehearing en banc). See Util. Air Regulatory Grp. v. EPA, 134 S. Ct. 2427, 2444 (2014) (internal quotation marks omitted) (“We expect Congress to speak clearly if it wishes to assign to an agency decisions of vast economic and political significance.”); see also Loving v. United States, 517 U.S. 748, 771 (1996) (“Congress may not delegate the power to make laws and so may delegate no more than the authority to make policies and rules that implement its statutes.”).

163 In this way, the requirements of bicameralism and presentment serve as a sort of overarching embodiment of due process: if the Executive enforces something that has not been properly enacted into law via the prescribed method, then the prosecution cannot be said to comport with the “due process of law,” even if the process itself is considered fair or adequate. See James W. Ely, Jr., “Due Process Clause,” The Heritage Guide to the Constitution, Heritage.org (“There are certain respects in which ‘due process of law,’ . . . uncontroversially regulates the substance of governmental action. . . . by ensuring that executive and judicial deprivations are grounded in valid legal authority. In this respect, the Fifth Amendment’s Due Process Clause limits the substance of executive or judicial action by requiring it to be grounded in law.”).

164 See The Federalist No. 62, at 376 (James Madison) (Clinton Rossiter ed., 1961) (observing that the “complicated check on legislation” imposed by the structural constraints in the Constitution “may in some instances be injurious as well as beneficial,” but defending it nevertheless as a safeguard against “the diseases to which our governments are most liable”).

FDA Poised to Spoil a Food Fight, Naturally

By Gregory D. Cote

Note from the Editor:

This article argues that the FDA should define the term “natural” because of the proliferation of litigation over alleged misuse of the term on food and beverage products and the fact that courts are ill-equipped to provide a definition.

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.


• Baylen Linnekin, The FDA’s Push to Define ‘Natural’ Food is an Exercise in Futility, REASON (May 14, 2016), http://reason.com/archives/2016/05/14/the-fdas-push-to-define-natural-food-is.


• Kelsey Albright, Should the word “natural” be banned from food labels?, Nat’l Consumers League (July 2014), http://www.nclnet.org/should-the-word-natural-be-banned-from-food-labels.

• Jason Best, The Case for Banning ‘Natural’ From Food Labels, TAKE PART (May 10, 2016), http://www.takepart.com/article/2016/05/10/natural.


About the Author:

Gregory D. Cote is a partner in McCarter & English’s Business Litigation practice group and represents companies in the food and beverage sector in a variety of food-related matters. He also specializes in constitutional and business litigation and has appeared in state and federal courts throughout the country.

World-renowned chef Thomas Keller once said, “Food should be fun.” But for those who make many of the products we eat and drink, there is nothing fun about food and beverage labeling litigation. These lawsuits have scorchéd federal dockets and reached gluttonous proportions over the last decade—rising from approximately 20 active cases in 2008 to nearly 500 as of this writing. The cases all follow the same standard recipe: a putative class action alleges that the use of certain marketing terms on a product’s label or packaging violates a state consumer protection statute because the terms purportedly provide false or misleading information about the product’s ingredients or nutritional attributes.

Amid this burgeoning tempest in a teapot, the U.S. Food and Drug Administration (FDA) repeatedly declined to weigh in on the definition of the term “natural,” which serves as one of the primary leaveners for these lawsuits. Instead, the agency opted to proffer only an informal guidance on the term.1 Egged on by what they saw as a gap in the FDA’s regulatory oversight, consumer and health advocacy groups filed the initial labeling lawsuits to prune what they described as the “health halo” effect of food and beverage labels.2 Soon thereafter, plaintiffs’ lawyers stuck their fingers in the pie as well and whipped up batches of other claims targeting everything from soup to nuts. Their theories and claims have evolved over the years, and now food and beverage makers are routinely embroiled in litigation over the presence of the word “natural” on their labels, even where the description manifestly describes the products at issue.3 Moreover, the FDA’s guidance interprets the term “natural” to mean that “nothing artificial or synthetic (including all color additives regardless of source) has been included in, or has been added to, a food that would not normally be expected to be in the food.” Food Labeling: Nutrient Content Claims, General Principles, Petitions, Definitions of Terms; Definitions of Nutrient Content Claims for the Fat, Fatty Acid, and Cholesterol Content of Food, 58 Fed. Reg. 2302 at 2407 (Jan. 6, 1993). The FDA also stated, however, that its guidance did not “address food production methods, such as the use of genetic engineering or other forms of genetic modification, the use of pesticides, or the use of specific animal husbandry practices, nor did it explicitly address food processing or manufacturing methods, such as thermal technologies, pasteurization, or irradiation. Furthermore, [the FDA] did not consider whether the term ‘natural’ should describe any nutritional or other health benefit.” Use of the Term “Natural” in the Labeling of Human Food Products, Request for Information and Comments, 80 Fed. Reg. 69905 at 69906 (Nov. 12, 2015); see also FDA, “Natural” on Food Labeling, available at www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/LabelingNutrition/ucm456090.htm (last visited April 25, 2018).


For example, a lawsuit against Sargento Foods, Inc. alleges that the company’s cheese is improperly labeled as “natural” because it is made with milk derived from cows that consumed genetically modified feed and treated with recombinant bovine growth hormone. However, the company’s cheese does not contain any unnatural ingredients, and the
some food and beverage makers, presumably looking to carve out their own piece of the pie, have filed labeling lawsuits against their competitors following the U.S. Supreme Court’s decision in *POM Wonderful LLC v. Coca-Cola Co.* 4 It’s a fine kettle of fish to say the least and a “natural” fit for class litigation.

The problem, in a nutshell, is that the meaning of the term “natural” varies by context in the food and beverage industry—what it means for a bag of chips or a can of ginger ale is different than for a package of cheese, a bottle of vegetable oil, or fresh produce. How food and beverage makers use the term matters as well—the phrase “Made With Natural Ingredients” generally means something different than “100% Natural” or “All Natural.” But context is not baked into these cases. Rather, the cases arbitrarily equate the term “natural” with purity and effectively disregard production and processing methods and the functional purpose of some ingredients. The most frequently targeted products are those made with ingredients that are genetically modified or sourced from animals that consume genetically modified feed, ingredients that allegedly contain incidental remnants of processing, and products with incidental additives used for flavoring, color, or other functional purposes. 5

Both large and small companies—including start-ups and family-owned companies—are targeted in the various types of labeling lawsuits. 6 If not dismissed at the outset, a “natural” lawsuit can put these companies in quite a pickle because they must either settle early or spend a lot of dough to defend it, risking harm to their brand and liability for damages and attorney fees, which can add up to tens of millions of dollars and bankrupt some companies. Litigating also risks establishing an unfavorably subjective judicial definition of “natural” rather than a uniform regulatory definition based on FDA deliberation, experience, and expertise. That is a bitter pill for most companies to swallow, so these cases are usually settled regardless of their merit (or lack thereof). Unfortunately, the settlements do little, if anything, to clarify the meaning of “natural” and are frequently just gravy trains for the plaintiffs’ attorneys, who obtain exorbitant fees for themselves, but gather merely crumbs—a nominal cash payment or a coupon or voucher—for the consumers they purported to represent. It’s a rotten outcome and leaves both consumers and food and beverage makers with a bad taste in their mouths. 7

But the FDA is now in the mix and could soon serve up a more appetizing option. After previously declining to define “natural” in the 1990s and subsequently declining the invitation of at least three federal courts to do so in 2013, the FDA announced the opening of a pre-rulemaking request for public comment regarding how it should define the term on November 12, 2015. 8 By the time the comment period closed on May 16, 2016, the FDA received more than 7,600 comments from individuals, consumer and industry groups, and food and beverage makers. Since then, the agency has moved slower than molasses in January and still has not issued a “natural” definition.

The FDA’s November 12, 2015 announcement initially led to a dip in the number of new “natural” lawsuit filings and prompted many courts to enter stays in pending cases based on the primary jurisdiction doctrine. 9 But the FDA’s silence on the issue since the announcement seems to have induced a slight resurgence of lawsuit filings in 2017 and now, in response to arguments by the so-called Plaintiffs’ Food Bar, some courts are beginning to deny new stay requests and to lift stays previously entered. These lawyers argue that the courts are a more expedient and appropriate forum for resolving the “natural” issue because the FDA’s extended delay indicates that it either does not intend to define “natural” or has not yet determined how or to define the term.

However, the FDA’s top banana, Commissioner Scott Gottlieb, M.D., recently made comments that could leave these lawyers with egg on their faces. During his keynote address at the National Food Policy Conference on March 29, 2018, the Commissioner defined “natural” as “the complete absence of chemicals and the complete presence of nature.”

---

5. The food and beverage industry is not the only target of these “natural” lawsuits. They also target personal care products (like toothpaste, cosmetics, baby wipes, and lip balm), household cleaning products (like window cleaner, dishwasher detergent and laundry detergent), furniture polish, and even pet food.
7. Class action settlements that primarily benefit the plaintiffs’ lawyers are facing increased scrutiny in the courts as well. Although not a “natural” case, the rejection of the Subway footlong sub settlement, which the Seventh Circuit described as a “racket” because it resulted only in fees for the class counsel and provided no meaningful relief for the class members, is perhaps the most notable recent example. See In re Subway Footlong Sandwich Mktg. & Sales Practices Litig., 839 F.3d 551, 552, 556-57 (7th Cir. Aug. 25, 2017).
8. See *Use of the Term “Natural” in the Labeling of Human Food Products, Request for Information and Comments*, 80 Fed. Reg. 69905 (Nov. 12, 2015). Relatedly, although the FDA defined the term “healthy” in 1994, the term’s use on food and beverage labels has provided ample fodder for litigation as well. However, on September 28, 2016, the FDA recognized that its definition of “healthy” had exceeded its shelf life and opened a pre-rulemaking request for public comment regarding how it should update its definition of that term. See *Use of the Term “Healthy” in the Labeling of Human Food Products; Request for Information and Comments*, 81 Fed. Reg. 66562. The public comment period closed on April 26, 2017. The FDA is still in the process of reviewing the comments and considering how to revise the term’s definition.
9. Primary jurisdiction is a discretionary doctrine invoked by courts to stay or dismiss claims that are properly cognizable in court, but involve issues that fall within the specialized competence of an administrative agency. The doctrine furthers uniformity and consistency in the regulation of issues within an agency’s purview and promotes better informed legal rulings by enabling courts to defer to and rely on agency expertise regarding technical and policy-related issues. See Reiter v. Cooper, 507 U.S. 258, 268–69 (1993); Nader v. Allegheny Airlines, Inc., 426 U.S. 290, 303 (1976); Weinberger v. Benton Pharm, Inc., 412 U.S. 645, 652-54 (1973).
Commissioner Gottlieb outlined the FDA's new multi-year plan for improving public health and enabling Americans to make better nutritional choices. Per Commissioner Gottlieb, defining "natural" is a key component of the consumer information and labeling initiatives that are part of the FDA's plan. Although he did not offer any definitive information about how or when the agency will define "natural," he confirmed that the agency reviewed all of the comments received in response to the November 12, 2015 announcement and acknowledged that "consumers increasingly want to know what is in the food they eat . . . [and] are trusting in products labeled as 'natural' without clarity around the term."11 He also noted that "there are wide differences in beliefs regarding what criteria should apply for products termed 'natural,'" but said the agency believes that "the natural claim [on food and beverage labels] must be true and based in science" and promised that it will "have more to say on the issue soon."12

These comments echo statements previously attributed to Commissioner Gottlieb in *Wall Street Journal* and *New York Times* articles.13 All combined, these statements confirm that the FDA continues to ruminate on the "natural" issue and expects its efforts to bear fruit soon. Arguments to the contrary and criticism of the FDA by the Plaintiffs' Food Bar are, therefore, not only half-baked, but self-serving and erroneous.

Some legal commentators have also suggested that recent defense victories indicate that the judiciary is beginning to curtail the "natural" lawsuits notwithstanding the FDA's inaction.14 Although well-intended and great food for thought, this assertion seems a bit too pie in the sky and should not be taken to indicate that FDA action is not needed. True, some courts have recently dismissed lawsuits concerning products purportedly containing trace amounts of glyphosate as well as suits concerning cooking oil made with bioengineered corn and yogurt made with milk from cows that consumed genetically modified feed. But cases based on the same or very similar allegations were allowed to proceed in other courts, and it is questionable whether the grounds for some of the recent dismissals can withstand appellate scrutiny. Thus, it is best to take such prognostications with a grain of salt because the results in the courts are mostly mixed and lack consistency.

Moreover, these inconsistent rulings demonstrate precisely why judicial deference to the FDA on primary jurisdiction grounds is warranted. Although courts routinely adjudicate consumer protection claims, that alone does not cut the mustard for these cases because they are not about merely alleged consumer confusion. For courts to resolve claims that defendants falsely and deceptively labeled their products "natural," they must first determine whether the defendants used "natural" ingredients and production/processing methods. Such determinations require consideration of complex technical, scientific, and policy issues that lie outside the judicial ken, but squarely within the FDA's expertise, experience, and congressionally delegated regulatory authority.

Nor are courts well-suited for establishing a uniform nationwide definition of "natural" for food and beverage labeling purposes. Rather, they are functionally equipped to resolve discrete cases, based only on the evidence and arguments presented by the parties to each case. Adjudicating the meaning of "natural" on a case-by-case and product-by-product basis in nearly 500 separate lawsuits is a costly and slow process and will almost assuredly produce discordant and subjective rulings that impose a patchwork of labeling standards and requirements that vary by jurisdiction. These rulings may also conflict with the definition Commissioner Gottlieb said is forthcoming from the FDA, which risks improperly elevating enforcement of state consumer protection statutes above the FDA's authority to establish food labeling laws.

Congress bestowed on the FDA the discretion and authority to establish labeling standards and determine whether and how to define "natural" for food and beverage labeling purposes.15 If rendered independent of FDA input, judicial determinations as to whether products are properly categorized as "natural" intrude upon the FDA's discretion and authority, undercut its expertise, potentially inhibit the uniformity of its regulatory and labeling regime, and potentially burden food and beverage makers by subjecting them to conflicting labeling standards.16 This is exactly what the primary jurisdiction doctrine is designed to prevent.

In fairness to the FDA and its seemingly slow-roasted approach to the task, defining "natural" is no piece of cake. The diversity of opinions in the comments received by the FDA makes defining "natural" a hard nut to crack. Technological advances in how ingredients are grown, harvested, and processed add even more complexity to the task. Plus, the agency's plate is already full with implementing the Food Safety Modernization Act as well

---

10 The speech, entitled Reducing the Burden of Chronic Disease, indicated that the FDA intends to build upon, not roll back, certain aspects of the Obama administration's healthy eating agenda, contrary to concerns expressed about the Trump administration's anti-regulatory agenda by health and nutrition advocates. The text of the speech is available on the FDA's website at www.fda.gov/NewsEvents/Speeches/ucm603057.htm.

11 Id.

12 Id.


16 Several courts have rejected similar arguments seeking dismissal based on the related doctrine of implied preemption because the FDA has not yet taken regulatory action regarding use of "natural" that is entitled to preemptive effect. See Holk v. Snapple Beverage Corp., 575 F.3d 329, 339–42 (2009); see also Garcia v. Kashi Co., 43 F. Supp. 3d 1359, 1373–74 (S.D. Fla. 2014); In re Frito-Lay N. Am. All Nat. Litig., No. 12-MD-2413, 2013 WL 4647512 at *10 (E.D. N.Y. Aug. 29, 2013). Thus, courts should instead invoke the primary jurisdiction doctrine pending FDA's anticipated action, which is likely to have preemptive effect and provide clarifying guidance for resolving the "natural" lawsuits.
as the new requirements for the Nutrition Facts panel and menu labeling—not to mention navigating a change in presidential administrations. That the FDA is still chewing on the issue is understandable given the circumstances.

Although promises—like eggshells and pie crusts—are made to be broken, the FDA finally appears poised to provide clarity regarding use of the term “natural” on food and beverage labels. In light of Commissioner Gottlieb’s consistent and repeated messaging, food and beverage makers that are defending “natural” lawsuits should promptly request stays on primary jurisdiction grounds. The courts should grant such requests and stay all pending “natural” lawsuits to avoid inconsistent outcomes and conserve judicial and litigant resources until the conclusion of the FDA’s proceedings. However, given the ongoing risk of litigation and the inconsistent manner in which the courts have managed the “natural” lawsuits thus far, food and beverage makers not presently defending a “natural” lawsuit should consult with counsel to assess their potential exposure and take appropriate mitigation measures to avoid getting burned by one of these lawsuits.
Why DoD Should Adopt a Multi-Cloud IT Strategy

By Marcia G. Madsen, Peter O. Schmidt, Luke P. Levasseur, David F. Dowd

Note from the Editor:

This article describes how companies approach cloud based services, and it argues that, in keeping with statutes that mandate competitive procurements, the Department of Defense should follow the lead of the commercial marketplace by adopting a multi-cloud strategy.

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.


About the Author:

Marcia G. Madsen, Peter O. Schmidt, Luke P. Levasseur, and David F. Dowd all practice law at Mayer Brown LLP, which represents Oracle Corporation.

Technology changes at the speed of light. Twelve years ago, there was no such thing as an iPhone or android mobile device. When the iPhone was introduced in 2007, Blackberry held a dominant position in the mobile communications market; it has less than 0.1% of the market today. Ten years ago, there was no iPad, Alexa, Uber, Instagram, Snapchat, Kickstarter, or Square. And although the idea of networked computing traces its lineage back to the 1960s, the term “cloud computing” wasn’t coined until 2006—and it is just within the last ten years that commercial cloud-based services and storage offerings have exploded. In light of this pace of development, no reasonable consumer—or large IT buyer—would lock itself into a single technology or service as its exclusive choice for the next decade.

In early March 2018, the Department of Defense (DoD) issued a draft request for proposal (RFP) for the Joint Enterprise Defense Infrastructure (JEDI) contract. It issued an amended JEDI RFP on April 16, 2018, and a final solicitation may be issued later this summer. The JEDI procurement involves cloud computing infrastructure as part of DoD’s effort to modernize its IT services. JEDI will apply across DoD and is valued in the billions of dollars. JEDI is not an acquisition for a single, broad-ranging cloud project, but is for an indefinite delivery, indefinite quantity (IDIQ) contract vehicle, under which work will be parceled out through separate orders for different requirements (and at different security levels) as particular agency needs are identified.¹

The JEDI procurement has received substantial attention—and been the subject of vigorous debate—because of size of the contract (estimated at $10 billion) and the projected size of the government cloud services market. Virtually all knowledgeable analysts recognize that DoD will experience significant efficiency (and other) gains from migrating much of its IT services to cloud-based facilities.² One of the most intensely debated aspects of JEDI is whether DoD should migrate applications and storage to a single cloud service provider (CSP), or multiple CSPs.

DoD has asserted that it intends to award JEDI to a single vendor. As we explain below, relying on a single vendor for JEDI cloud services would be a serious and unnecessary error. As the commercial marketplace demonstrates, multiple cloud solutions reduce enterprise costs, increase agility, insulate customers from problems from a single point failure, and offer substantial performance and security benefits. No evidence-based or coherent explanation for selecting a single-cloud provider for JEDI has been provided, and doing so will stifle innovation and increase government costs.

¹ See FAR 16.504(a).

² See Phil Goldstein, DOD, State Department See Benefits from Shifting Global Operations to the Cloud, FEDTECH (July 14, 2017), https://fedtechmagazine.com/article/2017/07/dod-state-department-see-benefits-shifting-global-operations-cloud; and Joint Enterprise Defense Infrastructure (JEDI) Cloud DRAFT Statement of Objectives (SOO) (April 16, 2018), https://www.fbo.gov/utils/view/id-09e2b02eb15b49a0b4e37ad121d8dec5, at 1 (draft JEDI RFP, Statement of Objectives, explaining that, among other things, large-scale migration to a cloud is necessary to avoid “environments [that] are not optimized to support large, cross domain analysis using advanced capabilities such as machine learning and artificial intelligence to meet current, and future 17 warfighting needs and requirements”).
competitors have evolved an approach that involves pushing some control on a market can experience substantial rewards into the cloud computing market. The common platform sales pitch has now metastasized Gateway Routing Protocol and Apple’s products in its consumer concerns have arisen with respect to Cisco’s Enhanced Interior lines of the same company the path of least resistance. Such result in high exit barriers and make expanding into other product point-of-contact for customer support—walled-off ecosystems efficient way to implement IT solutions. Although a common adoption of a single company’s technology ecosystem as the most vehicle, that is a lot of eggs to place in one basket. For the flexibility of a competitive multi-vendor IDIQ contract change, the time and money at stake, and the statutory preference. This procurement has aroused much interest because industry understands that cloud computing is likely to develop into a massive and rapidly changing market. As with other technology fields, new entrants are appearing frequently, and leading-edge capabilities are changing rapidly. As a result, in mid-2018, it is not possible to predict which technology or CSP or approach will be the most capable in three years, or which company might provide the best pricing in two years. Given this changing environment, the government should put itself in a position to take maximum advantage of innovations in the market and not tie its hands. But DoD is ignoring the manner in which commercial buyers are reacting to the changing cloud services environment (and the recommendations of industry) and, instead, apparently intends to proceed with a single-vendor approach that will preclude any further consideration of options for a particular application (under JEDI).

It is possible that DoD will do well by locking itself into a single company’s technology. But given the pace of technological change, the time and money at stake, and the statutory preference for the flexibility of a competitive multi-vendor IDIQ contract vehicle, that is a lot of eggs to place in one basket.

I. Background of Cloud-Based Services

A. The Industry Push for Ecosystems

Certain companies within the IT industry have touted mass adoption of a single company’s technology ecosystem as the most efficient way to implement IT solutions. Although a common platform can increase certain efficiencies—such as having a single point-of-contact for customer support—walled-off ecosystems result in high exit barriers and make expanding into other product lines of the same company the path of least resistance. Such concerns have arisen with respect to Cisco’s Enhanced Interior Gateway Routing Protocol and Apple’s products in its consumer market. The common platform sales pitch has now metastasized into the cloud computing market.

A company that can garner sufficient critical mass to exercise some control on a market can experience substantial rewards by promoting an ecosystem solution. To counter this strategy, competitors have evolved an approach that involves pushing cross-platform integrations. Such integrations allow individual participants to develop a best-of-breed technology in one area and integrate with the products of other companies (which may be best-of-breed in other areas). The resulting hybrid solutions can outperform any single vendor’s ecosystem.

Although various participants in the cloud computing market have pushed the single ecosystem concept, the structure of the market is contrary to a single-cloud environment. In part, that results from the fact that, unlike earlier IT markets, the best cloud-based services have been built from the ground up with the intention of integration. That intention is fundamentally thwarted by a walled-off system that makes integration across platforms difficult or impossible.

B. What Is Cloud Computing?

Generally speaking, cloud computing is a business model for renting access to shared software and hardware over a remote network—usually the internet. Although there are many nuances to cloud offerings, they are generally separated into Software as a Service (SaaS), Infrastructure as a Service (IaaS), and Platform as a Service (PaaS). Most users are familiar with SaaS functions such as creating a Google document in a web browser. In this case, the software managing the document is not running locally on the user’s computer, but remotely on servers owned and operated by Google. Such cloud-based services can be most sharply contrasted with on-premises solutions in which the user controls both the IT hardware—including acquiring, locating, maintaining, servicing, and repairing that hardware—and the software running on the hardware.

With IaaS, a cloud provider offers virtual hardware. The customer interacts with the hardware over the internet just as it would with a physical (albeit remote) server. For example, the customer might install a new web server on its IaaS, and then use that software (housed on cloud-based hardware) in the same manner as software housed on hardware at the user’s facility. PaaS refers to services that offer software resources, such as database software or a development environment, that developers can use to build more customized applications.

In providing services to customers, a CSP scales its servers up or down to meet demand. A customer will interact with the CSP through a single interface, while the CSP may use multiple physical servers (typically shared with other customers) to deliver the required computing and storage power. With basic computing resources such as storage, processing power, and network bandwidth available in ever greater supply through cloud computing services, prices have been falling, increasing the number of customers able to access and afford large scale computing systems.

---

3 10 U.S.C. § 2304a(d).

4 A cloud ecosystem is a complex system of interdependent components engineered to work together to enable cloud services. See Definition: Cloud Services, TechTarget, https://searchcloudcomputing.techtarget.com/definition/cloud-ecosystem.

C. Business Benefits of Different Cloud Environments

Commercial businesses and governments can choose to migrate applications and storage to only a single cloud or to use multiple cloud service providers (multi-cloud). Multi-cloud offerings provide at least four important benefits.

First, multi-cloud arrangements limit vendor lock-in and thereby increase cost competition. It is relatively easy to build a solution that is independent of any specific provider (and can therefore operate in multiple clouds) if this need is considered during the design phase. Unfortunately, naïve customers often find themselves with provider-specific solutions and, as a result, have little leverage in future negotiations (because they are locked in). Although a customer may have received bulk discounts and other cost savings up front, the vendor lock results in the customer being unable to take advantage of price reductions that result from the constantly declining cost of technology and the relentless commodification of cloud services. By pursuing a multi-cloud strategy from the outset, customers set themselves up for long term savings.

Second, the use of multiple CSPs allows customers to be agile and ramp up or down usage of a provider based on factors such as features, performance, and cost. For example, no one CSP will be the best on every metric in every region. Low latency may be important for a particular application in a particular region. But waiting for a certain performance level to be achieved by a CSP in a given region can compromise effectiveness. The ability to use multiple providers greatly increases coverage of resources (wherever located). Accordingly, customers are able not only to achieve target performance as early as possible, but they can effectively leverage price competition when other providers later build equivalent capability.

Third, a multi-cloud approach insulates customers from catastrophic failures by a single provider. With one cloud, a systemic failure by the provider (such as an unpatched security vulnerability or a design flaw) can result in substantial parts of the customer’s cloud system becoming inoperable. For example, during Amazon Web Services’ (AWS) February 2017 S3 storage outage, connected systems failed as well. A company’s or government’s use of multiple CSPs introduces redundancy that limits any failures to an isolated component or subsystem. Every provider’s systems are subject to human error and failure. Multi-cloud environments reduce the likelihood of system-wide problems as a result of such errors and failures.

Fourth, using a range of providers enables customers to take the path of least resistance for migrating each application. For example, it is generally faster and cheaper to migrate an Oracle on-premises database to an Oracle cloud database, as opposed to another CSP’s cloud database. Assuming several providers would perform equivalently for a particular application, significant savings can be achieved by this approach.

Very few large businesses opt to use or pursue a single-cloud environment for their applications. Customers in the CSP market have found that the principal theoretical benefit of having all applications and storage on a single cloud—i.e., fewer consistency problems resulting from the data being consolidated in one place with a single architecture—are less than the substantial benefits of having best-of-breed technologies from different offerees.

D. Technical Benefits of Multi-Cloud Environments

1. Performance

One of the principal benefits of multi-cloud is the ability to use best-of-breed implementations regardless of which entity developed a given implementation. This is facilitated by microservice architectures, in which applications are structured as collections of independently deployable services that communicate with each other. The same interfaces that allow developers to integrate a provider’s services into their workflows can be used to interact with other CSPs’ services. Examples of this capability can be found in “serverless” compute, such as AWS Lambda or Azure Functions, in which workloads are triggered by “hypertext transfer protocol” (http) webhooks or callback functions (whether from another provider, IaaS compute, or a customer’s on-premises service). In addition, the speed of inter-cloud communication can be very fast because cloud service providers have high speed connections to other providers.

2. Security

In addition to encryption and identification protocols, an important way companies (and CSP customers) improve security in cloud environments is by segmenting workloads across services and providers. This works based on the shared security model of working with CSPs. In short, CSPs certify that they provide security at the standard applicable to the systems for which they are responsible (e.g., from the host server’s hardware up to operating system). The customer is then responsible for ensuring that its data is secure the rest of the way. Thus, as with on-premises solutions, the customer must maintain security domains and separation of access. This provides the opportunity to segment workloads across services and providers, which further secures operations.

Even businesses or governments with a single cloud provider generally segment workloads, e.g., across multiple accounts and locations. Such segmentation not only allows for complex workloads to scale and to be broken down into manageable parts, it also limits the impact of system failures, human mistakes, and security events. The use of multiple providers further limits such potential problems because the systems are further separated. For instance, different providers have different security models. Enforcing logical access control (i.e., the tools and protocols used for identification, authentication, authorization, and accountability) at the provider level therefore makes it harder to grant accidental or unintended access to services that exist on a separate provider.
3. Support for Different Environments

Although using multiple providers for similar functions creates some concern regarding added complexity, an entire industry has developed within the cloud marketplace in response to the need to support multi-cloud environments. Tools to address complexity range from very specific to very broad in scope and cover everything from cost tracking and analysis to infrastructure and software management. It is easy to find support for any of the major CSPs. And because these tools are often focused on a specific issue, in many cases they address that issue better than the native tools of one provider. This is apparent in many budgeting and cost control tools, in which the graphs and reports are extensive and extensible. Likewise, the tools, such as Terraform for managing infrastructure, often are easier to use and provide a more complete picture of a customer’s overall system as compared with native tools.

These kinds of tools are not generally required to help single-cloud environments, which can result in upfront savings and faster procurement. However, these savings can be offset by ecosystem lock-in and worse performance from the inability to use best-of-breed solutions. In this regard, single-cloud environments may be better for small, quick-and-dirty projects, but are generally inferior for large, complex, long-lifetime projects.

II. The Legal Standard Applicable to IDIQ Procurements

As explained above, DoD has issued the JEDI procurement as a single-award IDIQ. However, procurement law favors competition. That preference for competition extends to the award of IDIQ contracts. Dating back to the mid-1990s, there has been a statutory preference for the award of multiple IDIQ contracts rather than a single award.

The preference for multiple IDIQ contracts has been strengthened over time, as the law now prohibits a single award by the agency makes a mandatory formal determination that a single-award IDIQ (in which there will be no competition for orders) is advantageous for the government. Specifically, statutes applicable to DoD make clear that IDIQ contracts valued over $112 million may not be awarded to a single source unless the head of the agency determines in writing that:

(A) the task or delivery orders expected under the contract are so integrally related that only a single source can reasonably perform the work;

(B) the contract provides only for firm, fixed price task orders or delivery orders for—

(i) products for which unit prices are established in the contract; or

(ii) services for which prices are established in the contract for the specific tasks to be performed;

(C) only one source is qualified and capable of performing the work at a reasonable price to the government; or

(D) because of exceptional circumstances, it is necessary in the public interest to award the contract to a single source.

JEDI encompasses a variety of cloud services to be implemented through various tasks across DoD at various classification levels, including unclassified work. Given the prevalence of the commercial cloud computing market, it is simply not credible to assert that only one company could perform such work at a fair and reasonable price.

Federal Acquisition Regulation (FAR) Part 16 implements the preference for multiple awards. The regulation provides that the contracting officer must, “to the maximum extent practicable, give preference to multiple awards” of IDIQ contracts. It emphasizes that each awardee “need not be capable of performing every requirement as well as any other awardee under the contracts.” Thus, the FAR expressly contemplates that contract awardees will be variously situated in terms of capabilities.

The purpose of the multiple award preference is to enable the government to obtain the benefit of recurring competitions for work that cannot be specifically defined initially but can be identified sufficiently with respect to discrete orders. Multiple awards give the government significant leverage. Companies must compete first to ensure they can be among the awardee group. The government then conducts competitions among the awardee group for task orders that likely will involve variation in needs. Multiple awards provide contractors with diverse strengths. Moreover, the contractors are incentivized to provide excellent performance at ever-more competitive prices throughout the contract term. Even the contractor with the best proposal at the contract stage may not have the capability to propose on every potential order, and each contractor must stay on its toes and sharpen its pencil in the successive rounds of competition at the order level. In multiple-award IDIQs, agencies that manage the contract well should hold a strong hand of cards that improves over time.

---


8 See, e.g., 10 U.S.C. § 2304, et seq.

9 See Pub. L. 103-355, § 1004.


11 FAR 16.504(c)(1)(i).

12 FAR 16.504(c)(1)(ii)(A).
III. DoD’s Stated Rationales for Single-Cloud Are Not Supported

On May 14, DoD released a congressional report that advances a number of arguments in favor of a single CSP. DoD’s explanation for pursuing a single-cloud IDIQ does not address the legal requirements for such a vehicle, though it indicates that it will release an explanation addressing these requirements at a later date. The decision to use a single award was apparently made without legal analysis, which apparently is being developed after the fact. Instead, DoD attempts to provide a technical explanation for its decision. In doing so, DoD purports to take a cautious approach by proceeding with a single cloud. But its arguments regarding why a single-cloud solution is purportedly cautious are rooted in at least seven crucial misunderstandings of multi-cloud solutions.

First, DoD states that “[r]equiring multiple vendors to provide cloud capabilities to the global tactical edge would require investment from each vendor to scale up their capabilities, adding expense without commensurate increase in capabilities.” In other words, DoD is concerned about the possibility of paying numerous vendors to scale-up such that they can provide capabilities worldwide (to the tactical edge), only to obtain duplicative capabilities.

In the commercial marketplace, cloud buyers have found that the use of multiple vendors enables providers to scale up faster and at lower cost. A multi-cloud approach would allow DoD to use different vendors in different use cases or environments where they already are optimal—instead of waiting (and paying) for one vendor to scale up in every area necessary. Although it is probably correct that “no other industry sector matches the scale and diversity of DoD’s tactical edge needs,” DoD recognizes that “certain industry sectors like oil and gas and university research have motivated vendors to develop commercial capabilities that can, at least to some degree, provide cloud computing and storage resources in austere and connectivity deprived environments.”

Those capabilities already exist across multiple vendors, and adopting a single-provider approach forces that provider to duplicate capabilities it does not already have.

In addition, DoD appears to recognize the market imbalance that would be created by paying only one provider to develop capabilities everywhere (to the “global tactical edge”) and thus plans to include in “the JEDI Cloud contract . . . a requirement for the contractor to provide a detailed portability plan.” Such a requirement will not ameliorate the imbalance DoD’s plan would create. Realistically, after one vendor has built to the global tactical edge (and been paid to do so), it will be difficult, if not impossible, for other vendors to compete.

Second, DoD asserts that “[w]hile security of data within clouds is largely standard and automatic, managing security and data accessibility between clouds currently requires manual configuration and therefore introduces potential security vulnerabilities, reduces accessibility, and adds cost.” In fact, an important premise underlying this assertion is incorrect. The security of data within clouds is not automatic: even with a single CSP, the customer must configure the environment and use encryption to fully control its data. Improperly configured environments can expose data, as evidenced by the numerous cases with AWS S3 buckets in the news.

Other security-related concerns, such as access control, do not indicate that implementation of a single-cloud approach is any safer than a multi-cloud environment. Tools such as Cisco CloudCenter manage access control and encryption across cloud environments. Using such tools, there is little difference between managing security for a single cloud or multiple clouds. Moreover, as explained above, segmenting workloads across services and providers increases security. Assuming DoD “make[s] extensive use of containerization,” “data standards,” and “application programming interfaces which expose the data over secure, modern protocols,” as it asserts it plans to do, it should be well-positioned to take advantage of this opportunity.

Third, DoD states that “[m]aintaining inconsistent and nonstandardized infrastructures and platform environments across classification levels complicates development and distribution of software applications, potentially adding delays and costs.” In other words, DoD believes it will be easier to develop and distribute applications on a single cloud infrastructure/environment, as a single system will be easier for personnel to learn and use—and easier to secure across classification levels.

The commercial market’s experience, in which customers manage data across different infrastructures and platforms, cannot be reconciled with DoD’s statement that unnecessary delays and costs will be added by using multiple clouds. This rationale is also inconsistent with DoD’s portability requirements. DoD acknowledges it “must strive to make applications portable,” which means platform-independent security. That required portability is inconsistent with the notion that there are substantial benefits to requiring use of a single cloud.

With respect to security, CSPs certify that they provide security for the systems for which they are responsible. DoD is

---

14 Id. at 4.
15 Id. at 9.
16 Id. at 12.
17 Id. at 4.
18 Under a shared security model (see Section D.2), the customer must use encryption to fully control its data, whether in transit or at rest.
21 Combined Congressional Report 10, 12.
22 Id. at 4.
23 Id. at 10.
responsible for security the rest of the way. Assuming DoD adopts a set of standards for CSPs and a set of platform-independent standards for application and data security, managing a multi-cloud environment across classification levels should require a small amount of additional work that is outweighed by the benefits of a multi-cloud environment.

Fourth, DoD states that the “use of multiple clouds would inhibit pooling data in a single cloud (i.e., a ‘data lake’), limiting the effectiveness of machine learning and artificial intelligence (AI).” This appears to be DoD’s principal objection, and it is repeated throughout the report. DoD asserts that “[m]arket research also indicated that initial migration to a single cloud is consistent with industry best practice”; the only rationale it provides to support that assertion is its concern regarding data lakes.

Data lakes are large repositories of raw data in native formats that, with the appropriate storage and processing tools, can be queried by a user. Even if the development and use of data lakes were “best practice”—and they are not—a single cloud is not necessary for a data lake, as companies have stepped in to fill the need for multi-cloud implementations. For example, Cloudera offers software for running multi-cloud Hadoop data lakes.

Contrary to DoD’s assertion, data lakes are not considered a “best practice” within industry. Industry recognizes that although use of a data lake benefits IT in the short term (as IT no longer has to devote resources to understanding how information is used when it is dumped into the data lake), getting value out of the data remains the responsibility of the business end user—and without at least some semblance of information governance, the lake ends up being a collection of disconnected data pools or information silos all in one place (a “data swamp”). Organizations have found that data lakes are expensive and time-consuming to coordinate, build, and maintain. With AI services having rich data source integrations, the cost and challenges of data lakes can be addressed by avoiding data lakes and allowing the AI services to pull from the original data repositories.

Data lakes also pose substantial risk to security and access control that have not been solved and are not addressed by DoD. Many data lakes are being used for data whose privacy and regulatory requirements are likely to represent risk exposure, but without enforced procedures for placement of data in a lake (where security capabilities and technologies have not been fully developed), it is not clear how the security requirements can be satisfied (particularly if left to non-IT personnel). Moreover, DoD does not explain how it will address the risks associated with a single location of data—or how a single point of failure would better facilitate resiliency, as compared to a scenario with multiple such points.

Finally, maintaining a large amount of data in a single CSP’s managed database service poses a high risk of lock-in. As one industry observer noted: “[M]any cloud vendors make it very difficult to extract data, configuration artifacts, and key application settings. This means that if rates rise, your freedom of movement is restricted. Even though your data is technically yours, it’s under the control and influence of someone else.” DoD does not explain how its vision of a data lake with a single CSP avoids the risk of potential lock-in.

Fifth, DoD notes that its experience to date shows that “hundreds of cloud initiatives have created numerous seams, incongruent baselines and additional layers of complexity for managing data and services at an enterprise level.” The report then asserts that “[s]cattering DoD’s data across a multitude of clouds further inhibits the ability to access and analyze critical data.” DoD’s concern is understandable. Interoperability is important for many reasons, including efficiency and maximizing data value. But while single-cloud solutions facilitate interoperability in some ways, by far the largest determinants of success with cloud migration are a business’ or government agency’s internal migration and development strategies. Thus, DoD’s experience to date does not represent a failure of multi-cloud; rather, it reflects the lack of a market research, planning, and a consistent strategy for cloud migration and management.

Sixth, DoD argues that it will capture some of the benefits of multi-cloud by contract, stating that “the JEDI Cloud contract will require ongoing commercial parity of technical offerings” and that “contract clauses [will] ensure DoD continues to get the best pricing as global marketplace pressures drive prices down.” Certainly, such clauses are better than nothing, and will

24 Id. at 4.
25 E.g., id. at 5 (“Leveraging ML/AI at a tempo required to be relevant to warfighters, however, requires significant computing and data storage in a common environment.”); id. at 6 (“The lack of a common environment for computing and data storage also will limit the effectiveness of ML/AI for warfighters.”); id. at 9–10 (“In addition to having a consolidated data lake, market research makes clear that a well-articulated data strategy, including an architecture and data storage standards, is critical to realizing the benefits particularly with regards to ML and AI.”).
26 Id. at 9. The report also quotes two Gartner reports, but those quotations only say that the transformation to the cloud should be staged in some way.
32 Id.
33 Id.
34 See id. at 7 (“The DoD’s adoption of cloud services to date has been mainly decentralized . . . .”).
likely provide some measure of savings. However, they are not as effective as a multi-cloud approach. Intellectual property laws and vigorous competition ensure that there is never true commercial parity of technical offerings—no software provider has ever been best-in-class in every area. And although most favored nation clauses (MFNs) can help offset the lack of bargaining power with a single CSP, DoD will be subject to the nuances of the terms, and the vendor will likely have an information advantage. For example, if an MFN clause covers a specific region but no other vendor operates in that region, the clause will have no effect and provide no benefit to the government.

Seventh, DoD acknowledges that “[i]f the commercial cloud marketplace offerings evolve to become interoperable and seamlessly integrated, DoD could have the ability to meet warfighting and business requirements by employing a range of future contract and award types.” This is not so much an independent rationale for a single-cloud approach as a restatement that one of DoD’s primary concerns is interoperability. But the evolution DoD is waiting for already has occurred and, although it is true that single-cloud solutions provide a modest benefit in facilitating interoperability, the best cloud technology has been built from the ground up with integration in mind. With a well thought out strategy and approach for cloud migration, management, and development, and with the use of currently available tools, DoD’s interoperability concerns can be fully addressed.

IV. Concerns Related to DoD’s Additional Reasons for Its Single-Award Strategy

DoD’s congressional report also attempts to defend offering JEDI as a single-award IDIQ by noting that, as currently structured, JEDI “only” calls for the award of a two-year base period, with the remainder of the ten-year term structured as options. Therefore, the government is not locked into a ten-year contract. That is true so far as it goes, but options typically are exercised where there is satisfactory performance by awardees. One thus reasonably should expect the options will be exercised for JEDI.

DoD also argues that JEDI will not be the only source for cloud services, as DoD already has stated in addressing the single-award approach. Although DoD has other vehicles under which it can procure cloud services, it surely intends to place heavy reliance on the JEDI vehicle. There is simply no reason to compete and award such a large contract if DoD does not intend to use it as a resource. Moreover, the argument that DoD will rely on multiple resources (including but not limited to JEDI) is inconsistent with the contention that DoD will be better off with a single JEDI provider.

Finally, the currently planned JEDI has been defended by people arguing that the contract will allow for new services to be added. Although that is correct, multiple-award IDIQs could similarly allow for the addition of other services, and competition would give incentive to add such services at the best possible price. It is not clear why the government has opted to forego the leverage it enjoys with multiple awards.

V. Conclusion

Today, most businesses already take a multi-cloud approach. This enables them to keep costs down, be agile, and insulate themselves from single point failure. It also provides them with enhanced performance (e.g., by combining best-of-breed services) and enhanced security (e.g., by segmenting workloads across multiple providers and enforcing logical access at the provider level).

As explained above, the relevant procurement statutes and the FAR make clear that Congress’ strong desire for competition in federal contracting extends to a preference for multiple-award IDIQ contracts whenever possible. Upon examination, the justifications for ignoring that preference and awarding a single IDIQ JEDI contract are not well supported. Best industry practices and the law counsel that DoD should consider this issue more carefully, follow congressional intent, and adopt a multiple-award approach for a large cloud procurement.
PARTY LIKE IT’S 1935?:
GUNDY v. UNITED STATES AND THE FUTURE OF THE NON-DELEGATION DOCTRINE

By Matthew Cavedon & Jonathan Skrmetti

Note from the Editor:

This article discusses Gundy v. United States, a case involving the Non-Delegation Doctrine in which the Supreme Court granted certiorari.

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.

• Ian Millhiser, How a strange Supreme Court case involving sex offenders could gut the EPA, Think Progress (March 5, 2018), https://thinkprogress.org/gorsuch-environmental-regulation-epa-892f82f3bc3/.

I. A Very Brief Overview of the Non-Delegation Doctrine

Article I of the Constitution famously limits the scope of federal legislative power to the items enumerated therein, and it vests that legislative power unequivocally in Congress.1 As the nineteenth century drew to a close, the Supreme Court could uncontroversially assert: “That congress cannot delegate legislative power to any other department of the government, executive or judicial, is an axiom of [the] legislative power can be delegated by congress to any other department of the government, executive or judicial, is an axiom ordained by the constitution.”2

The dissent in Marshall Field & Co. v. Clark, 143 U.S. 649, 692 (1892) (per Harlan, J.), 3 5 3 U.S.C. § 20913(d) (“The Attorney General shall have the authority to specify the applicability of the requirements of this title to sex offenders convicted before the enactment of this Act . . . .”).

1 34 U.S.C. § 20913(d) (“The Attorney General shall have the authority to specify the applicability of the requirements of this title to sex offenders convicted before the enactment of this Act . . . .”).
4 See U.S. Const. art. 1 § 1 (“All legislative Powers herein granted shall be vested in a Congress of the United States . . . .”).
5 Marshall Field & Co. v. Clark, 143 U.S. 649, 692 (1892) (per Harlan, J.). The dissent in Marshall Field—apparently a concurrence in judgment in modern parlance—agreed almost verbatim, stating “[t]hat no part of [the] legislative power can be delegated by congress to any other department of the government, executive or judicial, is an axiom in constitutional law, and is universally recognized as a principle essential to the integrity and maintenance of the system of government

About the Authors:

The views stated here are the authors’ alone and should not be attributed to the Federalist Society or their respective employers. Matthew Cavedon is a member of the Federalist Society’s Criminal Law & Procedure Practice Group Executive Committee. He is an Assistant Public Defender of the Northeastern Judicial Circuit in Gainesville, Georgia. Jonathan Skrmetti is a member of the Federalist Society’s Administrative Law & Regulation Practice Group Executive Committee. He is an attorney focusing on investigations, white collar defense, and complex litigation with Butler Snow, LLP in Memphis, Tennessee. He extends special thanks to Zachary Cookton for his able and enthusiastic assistance.

Herman Avery Gundy was convicted of violating the federal Sex Offender Registration and Notification Act (SORNA). Because Mr. Gundy’s underlying sex offense happened before SORNA’s enactment, the statute empowered the Attorney General to determine whether or not the Act applied to him.1 Mr. Gundy challenged this grant of authority by Congress to the executive branch, raising the Non-Delegation Doctrine. The U.S. Supreme Court agreed to hear his case, and oral arguments are scheduled for October 2, 2018.2

This grant of certiorari is remarkable, as many observers have thought (some approvingly, some lamentingly) that the Non-Delegation Doctrine was a non-starter this side of the Great Depression. A revival of the Doctrine could, depending on its nature, substantially limit the autonomy of the modern federal administrative state—a prospect that excites many conservatives and libertarians. On the other hand, a Supreme Court holding in Mr. Gundy’s favor could drastically shorten SORNA’s reach, which some right-of-center readers might deplore. And should the government prevail, the Court may effectively replace the principle of separation of powers with a new commitment to efficient technocracy, constitutionalizing the long trajectory of agency evolution at the expense of the Constitution’s original meaning. Either way, Gundy is likely to be a landmark decision.

This article will familiarize readers with the two closely-related legal issues at stake in Gundy—the Non-Delegation Doctrine as a whole, and whether that Doctrine is especially stringent when criminal penalties are on the line. The discussion to follow will provide the background needed to see how Gundy could be decided—and just how historically important it could be.

1 34 U.S.C. § 20913(d) (“The Attorney General shall have the authority to specify the applicability of the requirements of this title to sex offenders convicted before the enactment of this Act . . . .”).
4 See U.S. Const. art. 1 § 1 (“All legislative Powers herein granted shall be vested in a Congress of the United States . . . .”).
expedience eventually prompted a doctrinal shift to a more pragmatic theory of delegation. The Taft Court delineated Congress’ authority to delegate the legislative power to the executive in a unanimous opinion affirming the President’s adjustment of customs duties on the importation of barium dioxyde: “If Congress shall lay down by legislative act an intelligible principle to which the person or body authorized to fix such rates is directed to conform, such legislative action is not a forbidden delegation of legislative power.” This became known as the intelligible-principle test.

The Court revisited the delegation of legislative power in two landmark 1935 cases arising from the National Industrial Recovery Act. In the first, the Court recognized that “there are limits of delegation which there is no constitutional authority to transcend,” and held that a delegation of legislative power that established no policy, standard, or rule to constrain the delegation lacked an intelligible principle and thus failed to surmount the constitutional threshold. In the second, again addressing a section of the National Industrial Recovery Act, the Court held unconstitutional a sweeping and unbounded delegation of legislative power to the President to promote via regulation “fair competition.”

The tide of global war washed away the vestiges of the Great Depression, a switch in time saved nine, and the Non-Delegation Doctrine fell from cherished principle to inert theory. In 1948, the Court decided the Lichter case, in which it confidently upheld a statute authorizing the Maritime Commission to prevent “excessive profits” in certain government contracts. Lichter enumerated a variety of delegations that had been endorsed by the Taft and Hughes Courts as examples of Congress providing sufficient direction to the executive branch:

- ‘Just and reasonable’ rates for sales of natural gas, Federal Power Comm’n v. Hope Gas Co., 320 U.S. 591; ‘public interest, convenience, or necessity’ in establishing rules and regulations under the Federal Communications Act, . . . National Broadcasting Co. v. United States, 319 U.S. 190; prices yielding a ‘fair return’ or the ‘fair value’ of property, ordained by the constitution.” 143 U.S. at 697 (Lamar, J., dissenting). Perhaps anticipating the arc of the delegation cases to come, Justice Lamar critiqued the majority’s endorsement of what he viewed as an unconstitutional delegation of power to the President to regulate 

Between the New Deal and the new millennium, the Non-Delegation Doctrine flickered on in the occasional dissent or concurrence. In 1963’s Arizona v. California, an original jurisdiction case relating to water rights, Justice John Marshall Harlan, dissenting in part on behalf of himself and Justices William O. Douglas and Potter Stewart, objected to a congressional delegation that gave the Secretary of the Interior authority to allocate approximately 1.5 million acre-feet of water per year without any guidance as to how that allocation should be made. Justice Harlan was concerned that “[t]he delegation of such unrestrained authority to an executive official raises . . . the gravest constitutional doubts.” He identified “two primary functions vital to preserving the separation of powers required by the Constitution” served by the intelligible-principle test: “it insures that the fundamental policy decisions in our society will be made not by an appointed official but by the body immediately responsible to the people,” and it “provide[s] the courts with some measure against which to judge the official action that has been challenged.”

Seventeen years later, in Industrial Union Department, AFL-CIO v. American Petroleum Institute, the third case in the string cite to arise before Levy Leasing Co. at 431. Schechter Poultry distinguished Keppel & Bros. and noted that the scope of the “unfair methods of competition” language in the FTCAct addressed in Keppel & Bros. was “left to judicial determination as controversies arise,” whereas the operative language of the NIRA then before the Court “delegate[d] legislative power to the President to exercise an unfettered discretion to make whatever laws he thinks may be needed or advisable.” 291 U.S. at 304; Tagg Bros. also predated the 1935 non-delegation cases and addressed the authority of the Secretary of Agriculture to set rates for interstate commerce of livestock at a stockyard. See 280 U.S. at 431. Levy Leasing Co., the third case in the string cite to arise before non-delegation’s watermark year, addressed a vagueness challenge to state statutory language unrelated to any agency delegation, and was resolved by the Court through direct analogy to the Fifth Amendment’s “just compensation” standard. 258 U.S. at 249–50.

ordained by the constitution.” 143 U.S. at 697 (Lamar, J., dissenting). Perhaps anticipating the arc of the delegation cases to come, Justice Lamar critiqued the majority’s endorsement of what he viewed as an unconstitutional delegation of power to the President to regulate international commerce. Id. at 699.

6 Cf. Perez v. Mortgage Bankers Ass’n, 135 S. Ct. 1199, 1215, 1223 n.6 (2015) (Thomas, J., concurring in the judgment). Recent scholarship has argued that outside of 1935, the Non-Delegation Doctrine never provided a meaningful check on the scope of executive lawmaking. See generally Keith E. Whittington & Jason Iuliano, The Myth of the Nondelegation Doctrine, 165 U. Pa. L. Rev. 379 (2017). Whether or not this empirical claim is true, the Court’s transparency in endorsing delegations of legislative power to the executive branch steadily increased through the twentieth century.

7 J.W. Hampton, Jr., & Co. v. United States, 276 U.S. 394, 409 (1928) (per Taft, C.J.).


10 See Lichter v. United States, 334 U.S. 742, 785 (1948).

11 334 U.S. at 786 (internal citations abbreviated). Keppel & Bros. predated Schechter Poultry and Panama Refining and essentially applied a form of proto-Skidmore deference to the Federal Trade Commission’s interpretation of “unfair method of competition.” See 291 U.S. at 314. Schechter Poultry distinguished Keppel & Bros. and noted that the scope of the “unfair methods of competition” language in the FTCAct addressed in Keppel & Bros. was “left to judicial determination as controversies arise,” whereas the operative language of the NIRA then before the Court “delegate[d] legislative power to the President to exercise an unfettered discretion to make whatever laws he thinks may be needed or advisable.” 291 U.S. at 304; Tagg Bros. also predated the 1935 non-delegation cases and addressed the authority of the Secretary of Agriculture to set rates for interstate commerce of livestock at a stockyard. See 280 U.S. at 431. Levy Leasing Co., the third case in the string cite to arise before non-delegation’s watermark year, addressed a vagueness challenge to state statutory language unrelated to any agency delegation, and was resolved by the Court through direct analogy to the Fifth Amendment’s “just compensation” standard. 258 U.S. at 249–50.

12 See 373 U.S. 546, 603, 626 (1963) (Harlan, J., dissenting in part).

13 Id. at 626.

14 Id.

15 448 U.S. 607, 672–76 (1980).
lawmaking.\textsuperscript{16} Like Justice Harlan in \textit{Arizona}, Justice Rehnquist argued that robust enforcement of that standard would ensure that the most democratically responsive branch of government drove social policy, and that courts would have a meaningful metric to determine whether a delegation went too far.\textsuperscript{17} Although Justice Rehnquist enthusiastically endorsed application of the Non-Delegation Doctrine to invalidate unconstitutional delegations of legislative authority,\textsuperscript{18} he adhered to the pragmatic view that too robust a Non-Delegation Doctrine could frustrate the effectiveness of government, and he followed Chief Justice William Howard Taft’s admonition in \textit{J.W. Hampton, Jr., & Co. v. United States} that “delegations of legislative authority must be judged according to common sense and the inherent necessities of the governmental co-ordination.”\textsuperscript{19}

Despite these occasional outbreaks of non-delegation, the Court—including even the dissenters who endorsed the application of the Doctrine to invalidate certain exceptional delegations—continued to view the vast majority of congressional delegations to agencies as unremarkable and constitutionally sound.

The modern push for a more robust Non-Delegation Doctrine took shape when Justice Clarence Thomas penned a solo concurrence in \textit{Whitman v. American Trucking Associations}.\textsuperscript{20} Justice Thomas’ efforts differed substantially from the prior efforts to enforce the Non-Delegation Doctrine: rather than disputing whether Congress sufficiently articulated an intelligible principle, Justice Thomas took aim at the intelligible-principle test itself. As the majority fortified the existing case law built upon \textit{J.W. Hampton Jr. & Co.}, Justice Thomas endorsed a return to constitutional first principles at the expense of stare decisis.

Writing for the Court and endorsing the status quo, Justice Antonin Scalia emphasized that throughout its history, the Court:

> found the requisite “intelligible principle” lacking in only two statutes, one of which provided literally no guidance for the exercise of discretion, and the other of which conferred authority to regulate the entire economy on the basis of no more precise a standard than stimulating the economy by assuring “fair competition.”\textsuperscript{21}

His opinion echoed his earlier \textit{Mistretta} dissent in which he—while contesting a delegation of raw legislative power uncoupled from any exercise of executive or judicial power—conceded that “the scope of delegation is largely uncontrollable by the courts”;

\textsuperscript{16} Id.
\textsuperscript{17} See id. at 685–86.
\textsuperscript{18} Id. at 686.
\textsuperscript{19} Id. at 675 (quoting \textit{J.W. Hampton}, 276 U.S. at 351).
\textsuperscript{20} 531 U.S. 457 (2001). At the intermediate appellate level, a panel of the D.C. Circuit had held that the EPA’s interpretation of a statute, though not necessarily the text of the statute itself, violated the Non-Delegation Doctrine. \textit{Am. Trucking Ass’ns, Inc. v. U.S. E.P.A.}, 175 F.3d 1027, 1034 (D.C. Cir. 1999); \textit{see also Am. Trucking Ass’ns, Inc. v. U.S. E.P.A.}, 195 F.3d 4, 6–8 (1999) (modifying prior opinion on petition for panel rehearing).
\textsuperscript{21} Id. at 474. This passage, and the surrounding section, was joined by the entire Court except Justice John Paul Stevens. See 531 U.S. at 459.

\textsuperscript{22} \textit{Mistretta} v. United States, 488 U.S. 361, 416-17 (1989) (Scalia, J., dissenting) (discussed in more detail below).
\textsuperscript{23} Id. at 417 (quoted in part in \textit{Whitman}, 531 U.S. at 475). Justice Scalia touched on delegation of legislative powers again in \textit{City of Arlington v. F.C.C.}, 569 U.S. 290, 304 n.4 (2013), explaining that agency rulemaking and adjudication “take ‘legislative’ and ‘judicial’ forms, but they are exercises of—indeed, under our constitutional structure they must be exercises of—the ‘executive power.’” Chief Justice John Roberts, in dissent, agreed with this sentiment but pointedly noted that “the citizen confronting thousands of pages of regulations—promulgated by an agency directed by Congress to regulate, say, ‘in the public interest’—can perhaps be excused for thinking that it is the agency really doing the legislating,” \textit{City of Arlington}, 569 U.S. at 315 (Roberts, C.J., dissenting).
\textsuperscript{24} 531 U.S. at 487.
\textsuperscript{25} Id. In a separate opinion, Justice Stevens took a diametrically opposite position and rejected the notion that the Constitution places any limits on Congress’ ability to delegate legislative power. See 531 U.S. at 489 (Stevens, J., concurring in part and concurring in the judgment).
\textsuperscript{26} 135 S. Ct. 1225, 1240 (2015). The dispositive issue before the Court in \textit{Association of American Railroads} was whether Amtrak is a governmental entity. \textit{See id.} at 1228. The Court acknowledged that “questions implicating the Constitution’s structural separation of powers and the Appointments Clause” remained outstanding. \textit{Id.} Justice Samuel Alito wrote a separate concurrence addressing a number of constitutional issues and acknowledging, perhaps reluctantly, his understanding that “the formal reason why the Court does not enforce the nondelegation doctrine with more vigilance is that the other branches of Government have vested powers of their own that can be used in ways that resemble lawmaking.” 135 S. Ct. at 1237 (Alito, J., concurring).
at the intelligible-principle test and opined that “[a]n examination of the history of [legislative and executive] powers reveals how far our modern separation-of-powers jurisprudence has departed from the original meaning of the Constitution.”27 Justice Thomas traced “the idea that the Executive may not formulate generally applicable rules of private conduct” to the fundamental concept of the rule of law as it emerged in classical antiquity, then followed that line of thinking through Magna Charta to the Founding.28 His close analysis of the Court’s non-delegation jurisprudence intertwined with an appreciation of the principles underlying separation of powers as articulated by Locke and Hume, Coke and Blackstone, Madison and Hamilton, and living theorists such as Maurice Vile and Philip Hamburger.29

Marshaling the combined forces of law, reason, and prudence, Justice Thomas castigat ed the Court’s intelligible-principle jurisprudence as an abdication of the duty to enforce the separation-of-powers doctrine that defines the constitutional structure of the federal government. “To the extent that the ‘intelligible principle’ test was ever an adequate means of enforcing [the distinction between legislative and executive power], it has been decoupled from the historical understanding of [those] powers and thus does not keep executive ‘lawmaking’ within the bounds of inherent executive discretion.”30 The intelligible-principle test allowed Congress to provide such minimal guidance to the administrative state that executive branch recipients of the original meaning of the Constitution: The Government may create generally applicable rules of private conduct only through the proper exercise of legislative power.”32 Justice Thomas’ stirring opinion in Association of American Railroads has been cited in five opinions from the Courts of Appeals. Significantly, four of those were authored by Justice Thomas’ now-colleague, Justice Neil Gorsuch.33 Given the narrow window between publication of Association of American Railroads and Justice Gorsuch’s elevation to the Supreme Court—barely two years—one can reasonably infer that Justice Thomas’ position has achieved particular resonance with Justice Gorsuch. The exact extent of that influence should become clear when the Court resolves Gundy.

II. GUNDY BRINGS THE NON-DELEGATION DOCTRINE BEFORE THE SUPREME COURT

Herman Gundy was convicted of sexual offense in the second degree in Maryland in October 2005.34 He was sentenced to twenty years’ imprisonment, with ten years suspended. Committing that offense also violated the terms of Mr. Gundy’s federal supervised release arising from a prior federal conviction; he was sentenced to another twenty-four months’ imprisonment for that violation. Mr. Gundy served his state sentence and then his federal sentence, the latter first in Maryland and then in Pennsylvania.35 In 2012, as his federal sentence wound down, he was transferred to a halfway house in New York.36 He received a furlough to make that trip—from a federal correctional institute in Pennsylvania to the Bronx—unescorted; the terms of the furlough acknowledged that he remained in the custody of the Attorney General throughout his travels despite the lack of an escort.37 Following his stint at the halfway house, Mr. Gundy was released from custody to a residence in New York.38 Mr. Gundy did not register as a sex offender.39 The federal government indicted Mr. Gundy in January 2013 and charged him with violating SORNA.40

SORNA requires sex offenders, a category that includes Mr. Gundy, to register with the National Sex Offender Registry, and to update that registration whenever the sex offender travels in interstate or foreign commerce.41 SORNA applies prospectively to sex offenders convicted following its passage in 2006.42 For those convicted before 2006, such as Mr. Gundy, SORNA provides

---

27 Id. at 1240–41.
28 Id. at 1242–43.
29 Id. at 1242–45, 1252. Justice Thomas’ concern that delegation threatens the structural separation of powers, and thus undermines the interbranch checks and balances that serve to protect individual liberty, has been similarly embraced by key members of the Trump administration. See, e.g., Neomi Rao, Administrative Collusion: How Delegation Diminishes the Collective Congress, 90 N.Y.U. L. Rev. 1463 (2015) (arguing that delegation corrupts Congress’ collective interest in its institutional authority by incentivizing individual members of Congress to focus on influencing the executive branch agencies); Donald McGahn, 17th Annual Barbara K. Olson Memorial Lecture (Nov. 18, 2017). available at https://feddoc.org/conferences/2017-national-lawyers-convention/#agenda-item-barbara-k-olson-memorial-lecture.
30 Id. at 1250 (citing Whitman, 531 U.S. at 487 (Thomas, J., concurring)).
31 Id. at 1251.
32 Id. at 1252.
33 See Gutierrez-Brizuela v. Lynch, 834 F.3d 1142, 1154 n.5 (10th Cir. 2016) (Gorsuch, J., concurring); Caring Hearts Personal Home Services, Inc. v. Burwell, 824 F.3d 968, 969 (10th Cir. 2016); De Niz Robles v. Lynch, 803 F.3d 1165, 1171 (10th Cir. 2015); United States v. Nichols, 784 F.3d 666, 671 n.3, 672 n.4 (10th Cir. 2015) (Gorsuch, J., dissenting from denial of rehearing en banc) (discussed in more detail below). The fifth appellate citation to Justice Thomas’ opinion came from Judge Kent Jordan in Egan v. Delaware River Port Authority, 851 F.3d 263, 279 (3d Cir. 2017) (Jordan, J., concurring in the judgment).
35 Gundy, 804 F.3d at 143.
36 Id. at 144.
37 Id.
38 Id.
39 Id.
40 804 F.3d at 144.
42 804 F.3d at 142 (citing 42 U.S.C. § 16913(d)).
that its applicability is determined by the Attorney General.\textsuperscript{43} Exercising that delegated authority, the Attorney General passed regulations applying SORNA in full to pre-2006 offenders, with limited exceptions.\textsuperscript{44} The statute did not explicitly require this. Indeed, some have even said that Congress’ grant of authority was so amorphous that the Attorney General could have opted to exempt all sex offenders convicted prior to SORNA’s passage, cover all of them, or—as in fact happened—find some middle ground.\textsuperscript{45}

Without yet raising this argument, Mr. Gundy first moved to dismiss the indictment against him for failure to state an offense. He argued that he was required to register only after he had traveled to New York and thus could not have violated Section 2250(a)—the section defining the crime of failure to register—the elements of which must be satisfied sequentially.\textsuperscript{46} The district court granted the motion, rejecting the government’s argument that Mr. Gundy was required to register as soon as SORNA became retroactive by the Attorney General’s determination.\textsuperscript{47} The government appealed to the Second Circuit. The appellate panel reversed and held that SORNA’s registration requirements “attached [to Mr. Gundy] at the latest on August 1, 2008, the effective date of the Attorney General’s final guidelines.”\textsuperscript{48}

On remand, Mr. Gundy was convicted. He appealed to the Second Circuit. The court of appeals affirmed the district court, focusing its opinion on Mr. Gundy’s argument that his travel from Pennsylvania to New York did not trigger Section 2250(a). Nestled at the tail end of the opinion was a blanket rejection from Pennsylvania to New York did not trigger Section 2250(a).—the section defining the crime of failure to register—the elements of which must be satisfied sequentially.\textsuperscript{46} The district court granted the motion, rejecting the government’s argument that Mr. Gundy was required to register as soon as SORNA became retroactive by the Attorney General’s determination.\textsuperscript{47} The government appealed to the Second Circuit. The appellate panel reversed and held that SORNA’s registration requirements “attached [to Mr. Gundy] at the latest on August 1, 2008, the effective date of the Attorney General’s final guidelines.”\textsuperscript{48}

Mr. Gundy petitioned for certiorari, and the Supreme Court granted his petition with respect to his fourth question presented: “Whether SORNA’s delegation of authority to the Attorney General to issue regulations under 42 U.S.C. § 16913(d) violates antidelegation principles.”\textsuperscript{50} Mr. Gundy’s argument—foreclosed by United States v. Gazman,\textsuperscript{51} and made only for preservation purposes—that SORNA violates antidelegation principles.

Mr. Gundy’s brief is a muscular paean to the separation of powers, arguing that the Constitution prohibits delegation of legislative powers, particularly in the criminal context, and that the operative language of SORNA “impermissibly delegates quintessentially ‘legislative’ powers” and fails the intelligible-principle test.\textsuperscript{52}

A battalion of amici sallied into the case, all of them writing in support of Mr. Gundy’s cause.\textsuperscript{53} Many of the amici argue for a substantial change in the law to limit congressional delegations of legislative authority, which could be implemented via either a toothier intelligible-principle test or a more hermetic seal between the legislative and executive branches. Taking the opposite tack and recognizing that Gundy has the potential to mark a bold new era of non-delegation, the Araiza et al. brief instead argues for a restrained approach that simply applies the existing case law to find a lack of an intelligible principle in this case.

Alone against Mr. Gundy, the government filed a plucky brief arguing that SORNA as a whole supplies an intelligible principle to guide the Attorney General and limit the delegated authority.\textsuperscript{54} Because the mechanics of applying SORNA to previously convicted sex offenders created “practical problems,” the United States argues, Congress delegated power to the Attorney General to ensure that he had the flexibility to implement the requirements effectively.\textsuperscript{55} The crux of the argument is that:

The delegation of authority to address transition-period implementation issues concerning pre-Act offenders did not erase SORNA’s overriding objective to establish a comprehensive national system for the registration of sex offenders, designed to provide the broadest possible protection to the public. Congress merely delegated to the Attorney General the judgment whether that clear general policy would be offset, in the case of pre-SORNA sexual offenders, by problems of administration, notice and the like for this discrete group of offenders—problems well suited to the Attorney General’s on-the-ground assessment.\textsuperscript{56}

In addition, the government argued that “the limited scope of the authority SORNA confers on the Attorney General made detailed

\textsuperscript{43} Id.

\textsuperscript{44} See 804 F.3d at 142–43 (citing 72 Fed. Reg. 8894, 8896 (Feb. 28, 2007); 72 Fed. Reg. 30,210, 30212 (May 30, 2007); 73 Fed. Reg. 38,030, 38,063 (July 2, 2008); 75 Fed. Reg. 81,849, 81,850 (Dec. 29, 2010)). The regulations authorizing retroactive application of SORNA include a limited exception for sex offenders who “have been in the community for a greater amount of time than the registration period required by SORNA.” 73 Fed. Reg. at 38,046–47 (quoted at 804 F.3d at 143 n.3).

\textsuperscript{45} See, e.g., Nichols, 784 F.3d at 668-69 (Gorsuch, J., dissenting).

\textsuperscript{46} See Gundy, 804 F.3d at 144.

\textsuperscript{47} Id.

\textsuperscript{48} Id. at 145.

\textsuperscript{49} 591 F.3d 83, 91-93 (2d Cir. 2010).

\textsuperscript{50} United States v. Gundy, 695 Fed. App’x 639, 641 & n.2 (2d Cir. 2017).

\textsuperscript{51} 86 USLW 3438. The relevant statute was subsequently codified at 34 U.S.C. § 20913. See, e.g., Brief for Petitioner at 1, Gundy v. United States, No. 17-6086.

\textsuperscript{52} See Br. for Pet., Gundy v. United States (2018) (No. 17-6086).

\textsuperscript{53} The amici include the Cato Institute and Cause of Action; the Competitive Enterprise Institute, Reason Foundation, and Cascade Policy Institute; the National Association of Federal Defenders; Pacific Legal Foundation; Philip Hamburger’s New Civil Liberties Alliance; the Center for Constitutional Jurisprudence; William D. Araiza and 14 Other Constitutional, Criminal, and Administrative Law Professors; the American Civil Liberties Union; Scholars Whose Work Includes Sex Offense Studies; the Institute for Justice; the Downsize DC Foundation, DownsizeDC.org, the Free Speech Coalition, Free Speech Defense and Education Fund, U.S. Constitutional Rights Legal Defense Fund, Public Advocate of the United States, Gun Owners Foundation, Gun Owners of America, Conservative Legal Defense and Education Fund, and Restoring Liberty Action Committee; the National Association of Criminal Defense Lawyers; and the Becket Fund for Religious Liberty.


\textsuperscript{55} Id. at 25.

\textsuperscript{56} Id. at 28 (internal quotations, brackets, and citations omitted).
statutory direction unnecessary.” 57 Indeed, “the scope of the authority SORNA granted the Attorney General is no different for nondelegation purposes from the discretion to exempt otherwise covered individuals from the duty to register.” 58

The Court’s willingness to grant certiorari strictly on the non-delegation issue suggests that the government may have a challenging road ahead. Yet, as the broad spectrum of amici show, even if the Court decides to enforce the Non-Delegation Doctrine in this case, there are a variety of paths the Court could take that could have very different ramifications for administrative law and the separation of powers.

III. Five Options for the Court

Based on the Court’s jurisprudence, there are essentially five ways the Court could resolve Gundy. The first is to maintain the Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate the Attorney General to create a crime and then imprison people for violating it—such that an affirmance would obviate an affirmative endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate Scalian position that the intelligible-principle test is effectively non-justiciable based on the lack of a clear constitutional distinction between legislating and executive activity that looks like legislating. The second option would be to adopt the position Justice Stevens took in his Whitman concurrence and hold that the delegation of legislative power to the executive branch is altogether unremarkable and poses no constitutional problem. While it is possible the Court could either stay the course or adopt a broader endorsement of delegation, those seem like the least likely outcomes. The extraordinary nature of the delegation in Gundy—in which the challenged statute effectively allowed the Attorney General to create a crime and then imprison people for violating it—is such that an affirmance would obviate

The third option is to take up Chief Justice Rehnquist’s position, articulated in his American Petroleum Institute concurrence, that the intelligible-principle test should have teeth to “ensure[s] that Congress itself make[s] the critical policy decisions.” 59 Even dull and modest teeth would be enough, in the case of Gundy, to invalidate a delegation of authority that allows the executive branch to define a crime. This course would allow the Court to continue to constrain congressional abdication of legislative authority through the somewhat nebulous threat of a constitutional limit on delegations, but without deviating from its existing jurisprudence. Enforcement of the Non-Delegation Doctrine in Gundy could be seamlessly incorporated into the existing case law, and it would be identified in the future as the third example of Congress exceeding the outer limit of its authority. For Justices inclined to demark the boundary between executive and legislative powers while adhering to stare decisis, the path of Rehnquist may be the most palatable option.

The fourth option, the Thomastic option, is to reconsider whether the intelligible-principle test is the constitutionally appropriate means of limiting congressional delegations of lawmaking authority. Under this approach, the Court would determine that a strict separation of powers—implementing the Lockean and Madisonian concepts that ostensibly drove the Court’s nineteenth-century decisions—does not permit Congress to give executive agencies the authority to enact generally applicable rules of private conduct. Such a decision could be written with nuanced care to harmonize it with the prior line of cases, but it would nevertheless mark a substantial departure from current administrative law. 60 Should the Court elect to take this approach, it could be a profound development in constitutional law. Even if the Court charts a different course, it seems likely that this position will motivate memorable separate opinions from an enthusiastic subset of the Court.

The fifth option is an intriguing compromise: The Court could avoid making any broad pronouncements on the Non-Delegation Doctrine generally by bifurcating it between criminal and non-criminal cases. Such an approach could allow the Thomastics on the Court to apply a more robust concept of separation of powers to criminal cases while leaving the legal structure underlying the modern administrative state essentially undisturbed. The remainder of this article examines this fifth option in detail.

IV. Gundy Could Establish a Heightened Criminal Non-Delegation Standard

Perhaps, as a number of jurists and amici have suggested, there is an especially searching non-delegation inquiry when an administrative rule triggers criminal sanctions. This notion has been debated for over a century, and Justice Gorsuch is one key recent proponent.

A. The Long Debate Over Whether There Is a Heightened Non-Delegation Doctrine in Criminal Cases

Judges have disagreed about whether the threat of criminal punishment triggers greater scrutiny of delegations since at least the late 1800s. 61 It is axiomatic, from all the way back in English common law, that “A CRIME . . . is a [sic] act committed, or

57  Id. at 29.
58  Id. at 30.
59  448 U.S. at 687.
60  Justice Thomas has provided a helpful metaphor to explain his occasional willingness to reexamine precedent:

When you get a case, you have the last decision in the line. That's what's on your desk. . . . The last decision in the line is like a caboose on a train. Let's go from the caboose all the way up to the engine, and see what really went on, and let's think it all through. You might get up to the caboose and find out: Oh, there's nobody in the engine. . . . You say, "There's nobody driving the train. What happened? Where did we go wrong? Maybe we're headed in the wrong direction. Let's think it through."


omitted, in violation of a public law.” That has often been taken to mean that only a statute can criminalize an act. Indeed, it is why the Supreme Court abolished common-law crimes, holding that Congress alone “must first make an act a crime, [and] affix a punishment to it.”

Not long after common-law crimes’ abolition, the executive branch began criminalizing actions through statutory interpretations. But the Supreme Court resisted this trend. For instance, in the late 1800s, the federal government tried to prosecute an oleomargarine dealer for improper bookkeeping. A federal statute required oleomargarine manufacturers to keep proper books. But that provision did not facially apply to dealers, and the statute only criminalized dealers’ neglect of “the things required by law.” But the Treasury Department’s implementing regulations imposed bookkeeping requirements on dealers. In United States v. Eaton, the Supreme Court found the prosecution of the dealer unlawful. Its analysis was both statutory and based on the separation of powers. It began by observing that it had already abolished federal common-law crimes. After noting that “regulations [cannot] alter or amend a law,” the Court interpreted the statute to foreclose the prosecution. “It would be a very dangerous principle,” the Court insisted, “to hold that a thing prescribed by the Commissioner of Internal Revenue . . . could be considered as a thing ‘required by law’” pursuant to a statute. Again, a crime can only be established by a public law—that is, a statute. The Court established a clear-statement rule: “If Congress intended to make it an offence . . . it would have done so distinctly . . . .” While regulations may be lawful, “it does not follow that a thing required by them is a thing so required by law as to make the neglect to do the thing a criminal offence in a citizen, where a statute does not distinctly make the neglect in question a criminal offence.”

Despite Eaton, the next major relevant Supreme Court case upheld a prosecution. In United States v. Grimaud, there was “no act of Congress which, in express terms, declare[d] that it [would] be unlawful to graze sheep on a forest reserve.” Statutory law merely required—on pain of criminal sanctions—that citizens’ use of forest reserves “for ‘all proper and lawful purposes’” be subject to “the rules and regulations covering” such areas. The Court made quick work of a challenge to this:

If, after the passage of the act and the promulgation of the rule, the defendants drove and grazed their sheep upon the reserve, in violation of the regulations, they were making an unlawful use of the Government’s property. In doing so they thereby made themselves liable to the penalty imposed by Congress.

Grimaud distinguished Eaton, characterizing the Eaton prosecution as:

putting the regulations above the statute. . . . [W]hen Congress enacted that a certain sort of book should be kept, the Commissioner could not go further and require additional books; or, if he did make such regulation, there was no provision in the statute by which a failure to comply therewith could be punished.

This reading, while reasonable in itself, completely ignored Eaton’s focus on the separation of powers.

Eaton’s worries about criminalization by regulation briefly bobbed up in the Supreme Court’s only two cases sustaining non-delegation challenges. Panama Refining Co. v. Ryan connected Eaton-esque concerns to due process:

If the citizen is to be punished for the crime of violating a legislative order of an executive officer, or of a board or commission, due process of law requires that it shall appear that the order is within the authority of the officer, board or commission, and, if that authority depends on determinations of fact, those determinations must be shown.

There was also a whiff of special scrutiny for criminal matters in A.L.A. Schechter Poultry Corp. v. United States. The opinion noted that fair-competition codes carried criminal sanctions and coercive power. The Court disapproved that such power was being exercised “due to the effect of the executive action.” These side-notes did not go unnoticed. The Court held in Fabey v. Mallonee that both opinions “emphasized” the criminal features of the delegations at issue in subjecting them to particular

62 4 William Blackstone, Commentaries *5 (emphasis added).
64 United States v. Eaton, 144 U.S. 677, 686 (1892).
65 Id.
66 Id. at 685.
67 Id.
68 Id. at 687.
69 Id.
70 Id. at 688.
71 Id. at 687–88 (citing William Blackstone, 4 Commentaries *5).
72 Id.
73 Id. (emphasis added).
74 220 U.S. 506, 521 (1911).
75 Id.
76 Id.
77 Id. at 519.
78 Curiously, Grimaud did not distinguish Eaton in the most obvious way available: Grimaud concerned regulations governing the use of federal property, a core executive function, whereas Eaton addressed a regulation of private activity. See, e.g., Gubienso-Ortiz v. Kanuhele, 857 F.2d 1245, 1270 (9th Cir. 1988), vac’d & remanded by United States v. Chavez-Sanchez, 488 U.S. 1035 (1989) (Wiggins, J., dissenting) (acknowledging, but rejecting, such a “core functions” analysis).
79 293 U.S. 388.
80 295 U.S. 495.
81 Id. at 529.
82 Id. at 537–38.
non-delegation scrutiny. Based on *Fahey*, one might think there had emerged a clear rule in favor of tightening the delegation belt in criminal cases.

But other precedent from the same time period pointed in the other direction. In *M. Kraus & Bros., Inc. v. United States*, a case about price controls, the Court demanded nothing more than that the criminally prohibited conduct “be set forth with clarity in the regulations and orders which [the executive branch] was authorized by Congress to promulgate,” because “Congress has warned the public [through the statute] to look to that source alone to discover what conduct is evasive and hence likely to create criminal liability.” The *Kraus* Court followed *Grimaud* in distinguishing *Eaton* on statutory grounds. It did not mention any of the other cases discussed above.

For the next twenty years, concerns about delegating away the criminalization power were mostly limited to minority opinions. Dissenting in *Barenblatt v. United States*, Justice Hugo Black made a vagueness argument that echoed *Eaton*’s holding, emphasizing that statutes themselves have to clearly spell out what conduct is illegal: “[T]he standard of certainty required in criminal statutes is more exacting than in noncriminal statutes. This is simply because it would be unthinkable to convict a man for violating a law he could not understand.” Justice William Brennan picked up on this idea in his concurrence in *Eaton*. He did so while distinguishing delegations carrying criminal sanctions from other delegations. He fully granted that “Congress ordinarily may delegate power under broad standards,” but he thought that “[t]he area of permissible indefiniteness narrows . . . when the regulation invokes criminal sanctions and potentially affects fundamental rights.” Before the *Robel* defendant (a registered Communist employed at a shipyard) could be sent “to prison for holding employment at a certain type of facility,” Justice Brennan wanted to be satisfied that “Congress authorized the proscription as warranted and necessary.” Why? Because the public has a right to know whether power is flowing from the proper authorities:

[Persons engaged in arguably protected activity . . . must not be compelled to conform their behavior to commands, no matter how unambiguous, from delegated agents whose authority to issue the commands is unclear. The legislative directive must delineate the scope of the agent’s authority so that those affected by the agent’s commands may know that his command is within his authority and is not his own arbitrary fiat.]

The dual priorities voiced by Justices Black and Brennan—the public’s right to know what is expected of them and who it is that holds those expectations—apparently soon convinced their peers. *Smith v. Goguen* expressed dismay at “standardless” statutory language that “allows policemen, prosecutors, and juries to pursue their personal predilections.” While *Goguen* was a vagueness case, its holding could just as easily have come from *Schechter*: “Legislatures may not . . . abdicate their responsibilities for setting the standards of the criminal law.” This holding received more force from another vagueness case, *Kelender v. Lawson*, which reiterated that legislatures must tell people what they cannot lawfully do:

[T]he void-for-vagueness doctrine requires that a penal statute define the criminal offense with sufficient definiteness that ordinary people can understand what conduct is prohibited and in a manner that does not encourage arbitrary and discriminatory enforcement. Although the doctrine focuses both on actual notice to citizens and arbitrary enforcement, we have recognized recently that the more important aspect of the vagueness doctrine is not actual notice, but the other principal element of the doctrine—the requirement that a legislature establish minimal guidelines to govern law enforcement.” Where the legislature fails to provide such minimal guidelines, a criminal statute may permit “a standardless sweep [that] allows policemen, prosecutors, and juries to pursue their personal predilections.”

While criminal justice “is certainly a matter requiring the attention of all branches of government,” the legislature alone must clearly state the rules. The Court even went so far toward special standards for criminal cases as to find that its precedent “expressed greater tolerance of [vagueness in] enactments with

---

85 *M. Kraus & Bros., Inc.*, 327 U.S. at 620 n.4.
86 *But cf.* United States v. Evans, 333 U.S. 483, 486 (1948) (saying, with respect to an ambiguous statute, that “defining crimes and fixing penalties are legislative, not judicial, functions”).
87 To be sure, the due process Vagueness Doctrine is distinct from the Non-Delegation Doctrine. But the two have significant overlap. “A vague law impermissibly delegates basic policy matters to policemen, judges, and juries for resolution on an ad hoc and subjective basis, with the attendant dangers of arbitrary and discriminatory application”—concerns that also underpin the Non-Delegation Doctrine. *Grayned v. City of Rockford*, 408 U.S. 104, 108–09 (1972).
89 389 U.S. 258 (1967).
90 *Id.* at 274.
91 *Id.* at 275.
92 *Id.* at 272, 277.
93 *Id.* at 281 (internal citation omitted).
95 *Id.*
97 *Id.* at 361.
civil rather than criminal penalties because the consequences of imprecision are qualitatively less severe.98

But just as Grimaud had swiftly taken the wind out of Eaton’s sails nearly a century before, Mistretta v. United States did the same to this line of cases. This article has already discussed Mistretta, which upheld Congress’ establishment of the U.S. Sentencing Commission, and which featured a vigorous dissent by Justice Scalia.99 But he did not address whether criminal cases deserve special treatment; in fact, he expressed concern about a hypothetical “Medical Commission” deciding issues like “the withholding of life-support systems in federally funded hospitals, or the use of fetal tissue for research.”100 Still, the criminal character of the sanctions directly at issue obviously weighed on him. Following a lengthy discussion of just how much power the Commission had over “application of the ultimate governmental power short of capital punishment,” he concluded that “the basic policy decisions governing society are to be made by the Legislature.”101

Justice Scalia revisited this issue in two more dissents over two decades later. The first again decried what he took to be an unlawful delegation of criminalization power, without considering whether there is a heightened non-delegation standard:

[I]t is not entirely clear to me that Congress can constitutionally leave it to the Attorney General to decide—with no statutory standard whatever governing his discretion—whether a criminal statute will or will not apply to certain individuals. That seems to me sailing close to the wind with regard to the principle that legislative powers are nondelegable . . . .102

Later, Justice Scalia inched slightly closer to embracing a special standard, at least for statutory interpretation:

[L]egislatures, not executive officers, define crimes. . . . With deference to agency interpretations of statutory provisions to which criminal prohibitions are attached, federal administrators can in effect create (and uncreate) new crimes at will, so long as they do not roam beyond ambiguities that the laws contain. Undoubtedly Congress may make it a crime to violate a regulation, but it is quite a different matter for Congress to give agencies—let alone for us to presume that Congress gave agencies—power to resolve ambiguities in criminal legislation.103

Meanwhile, the Court as a whole hemmed and hawed. In Touby v. United States, decided in 1991, it candidly noted that its “cases are not entirely clear as to whether more specific guidance [than what the intelligible-principle test mandates] is in fact required.”104 In Loving v. United States, the Court seems to have come down firmly in favor of applying the ordinary non-delegation standard: “There is no absolute rule . . . against Congress’ delegation of authority to define criminal punishments.”105 A delegation is proper “so long as Congress makes the violation of regulations a criminal offense and fixes the punishment, and the regulations ‘confine[e] themselves within the field covered by the statute.’”106 But what about the public’s right to know by what authority criminal laws are promulgated? “The exercise of a delegated authority to define crimes may be sufficient in certain circumstances to supply the notice to defendants the Constitution requires.”107 Loving should be read cautiously, though, given its facts. It concerned whether or not the president had the power to identify aggravating factors in capital cases in courts martial.108 The majority specifically noted that this matter was within “the traditional authority of the President” over the armed forces.109 Loving is also limited by later precedent denying deference to administrative agencies’ interpretations of criminal statutes.110

Still, with sweeping proclamations both against and in favor of heightened non-delegation scrutiny, Touby remains correct that the “cases are not entirely clear.”111

B. Resolving the Constitutionality of Applying SORNA Retroactively Using a Heightened Non-Delegation Doctrine

Perhaps the right context for settling the dispute is SORNA’s retroactivity. Judge Carlos Lucero and then-Judge Gorsuch certainly thought so in 2015. Each dissented from the denial of

---

99 As discussed above, Justice Scalia would have held that the Commission was compatible with the Non-Delegation Doctrine. Mistretta, 488 U.S. at 416. His critical analysis, which shares some analytical overlap with the non-delegation precedent at issue here, was based instead on the Vesting Clause and the lack of any related exercise of executive or judicial power to provide constitutional cover. Id. at 416–17.
100 Id. at 422 (Scalia, J., dissenting).
101 Id. at 413, 415.
102 Reynolds v. United States, 565 U.S. 432, 450 (2012) (Scalia, J., dissenting). See also id. at 448 (deeming SORNA to be facially retroactive, and interpreting the supposed delegation at issue in Gundy “as conferring . . . an authority to make exceptions to the otherwise applicable registration requirements.”). Of course, if Gundy adopts this approach—in which Justice Ginsburg joined—then it can sidestep non-delegation. But the price would be overturning Reynolds, the outcome of which was supported by six current justices.
103 Whitman v. United States, 135 S. Ct. 352, 353 (2014) (Scalia, J., respecting denial of cert.) (internal citation omitted).
104 Touby v. United States, 500 U.S. 160, 166.
106 Id. (citing Grimaud, 220 U.S. at 518).
107 Id. (citing M. Kraus & Bros., Inc., 327 U.S. at 622).
108 Id. at 751–52.
109 Id. at 772; see also id. at 776 (Scalia, J., concurring in part and concurring in judgment) (“[I]t would be extraordinary simply to infer . . . a special limitation upon tasks given to the President as Commander in Chief, where his inherent powers are clearly extensive.”); id. at 778 (Thomas, J., concurring in judgment) (“By concurring in the judgment in this case, I take no position with respect to Congress’ power to delegate authority or otherwise alter the traditional separation of powers outside the military context.”). Perhaps this is the vindication of the ‘core functions’ approach mentioned above at footnote 78?
111 500 U.S. at 166.
rehearing en banc in United States v. Nichols. The panel decision in that case upheld SORNA’s retroactive application against a non-delegation challenge. With fairly little analysis, the panel “declined to abandon the well-settled ‘intelligible principle’ standard.” Judge Lucero concisely disagreed.

Judge Gorsuch provided a more thorough critique. His dissent was based on the separation of powers, with added support from concern for personal liberty. “If the separation of powers means anything,” he began, “it must mean that the prosecutor isn’t allowed to define the crimes he gets to enforce.” In fact, the power to criminalize is at the very heart of non-delegation:

Without doubt, the framers’ concerns about the delegation of legislative power had a great deal to do with the criminal law. The framers worried that placing the power to legislate, prosecute, and jail in the hands of the Executive would invite the sort of tyranny they experienced at the hands of a whimsical king.

As an antidote to such despotism, the Founders enshrined “the principle that the scope of individual liberty may be reduced only according to the deliberately difficult processes prescribed by the Constitution, a principle that may not be fully vindicated without the intervention of the courts.”

Judge Gorsuch reviewed much of the precedent discussed in this article. He found that “the [U.S. Supreme] Court has never expressly held that an intelligible principle alone suffices to save a putative delegation when the criminal law is involved.” What is more, that Court had never faced a situation as broad as the one in Nichols, where “legislation left it to the nation’s top prosecutor to specify whether and how a federal criminal law should be applied to a class of a half-million individuals.” Under the circumstances, Judge Gorsuch thought that it was “easy enough to see why a stricter rule would apply in the criminal arena”—liberty interests, originalism, and modern over-criminalization:

The criminal conviction and sentence represent the ultimate intrusions on personal liberty and carry with them the stigma of the community’s collective condemnation—something quite different than holding someone liable for a money judgment because he turns out to be the lowest cost avoider. Indeed, the law routinely demands clearer legislative direction in the criminal context than it does in the civil and it would hardly be odd to think it might do the same here. When it comes to legislative delegations we’ve seen, too, that the framers’ attention to the separation of powers was driven by a particular concern about individual liberty and even more especially by a fear of ending one set of hands with the power to create and enforce criminal sanctions. And might not that concern take on special prominence today, in an age when federal law contains so many crimes—and so many created by executive regulation—that scholars no longer try to keep count and actually debate their number?

According to Judge Gorsuch, perhaps no case better could have represented these problems than Nichols. “[T]he discretion conferred” was simply:

extraordinary—in its breadth (allowing the Attorney General to apply none, some, or all of SORNA’s requirements to none, some, or all past offenders), in its subject matter (effectively defining a new crime), in its chosen delegate (the nation’s top prosecutor), and in the number of people affected (half a million).

These considerations showed that “more, not less, guidance [wa]s required.”

For these reasons, and drawing inspiration from Toob, Judge Gorsuch would have adopted a three-part test for delegations implicating the power to criminalize. First, “Congress must set forth a clear and generally applicable rule”; second, that rule must “hinge[ ] on a factual determination by the Executive”; and third, “the statute provides criteria the Executive must employ when making its finding.” The delegation at issue in Nichols (and Gundy) easily failed that standard. Congress just “pointed to a problem that needed fixing and more or less told the Executive to go forth and figure it out.” Judge Gorsuch concluded: “By any plausible measure . . . that is a delegation run riot, a result inimical to the people’s liberty and our constitutional design.”

C. Justice Gorsuch’s View Evolved

But Justice Gorsuch’s reasoning from Nichols may not carry over to Gundy. Consider his partial concurrence in Sessions 2018 The Federalist Society Review 51
Dimaya concerned a vagueness challenge to the definition of “crime of violence” in the Immigration and Nationality Act. An alien found to have committed such a crime suffered civil penalties such that removal from the country became “a virtual certainty.” The Dimaya majority framed the decision in terms of vagueness. A plurality of the Court emphasized both the importance of immigrants having fair notice of what consequences they could face, and vagueness as “a corollary of the separation of powers—requiring that Congress, rather than the executive or judicial branch, define what conduct is sanctionable and what is not.” The plurality rejected the government’s contention that a lower vagueness standard should have applied in Dimaya than in criminal cases, observing that deportation is a severe consequence and that immigration is closely tied to criminal adjudications. Following criminal precedent regarding a similar statute, a majority of the Court found the definition of “crime of violence” to be unconstitutionally vague.

The Dimaya plurality did not decide whether there is a special non-delegation standard for criminal cases. Its refusal to weaken the criminal vagueness doctrine for a civil case might hint that the degree of scrutiny always depends on the severity of the real-world consequences. But Gundy is Dimaya’s procedural inverse. In Dimaya, the government asked the Court to lower its criminal vagueness standard for a civil case. In Gundy, the question is whether the non-delegation standard rises when criminal sanctions loom. That might be a distinction with a difference. The Dimaya plurality simply did not say enough to tell.

But Justice Gorsuch showed his cards. His hand has changed somewhat in the four years since his Nichols dissent, and he now thinks there is only one vagueness/non-delegation standard for both civil and criminal cases. He reached this conclusion through historical study and practical concerns. Examining early precedent, Justice Gorsuch found that “[c]ourts refused to apply vague laws in criminal cases involving relatively modest penalties—and “in civil cases too.” Justice Gorsuch then turned to modernity. Given that “the severity of the consequences counts when deciding the standard of review,” he invited the Court to “take account of the fact that today’s civil laws regularly impose penalties far more severe than those found in many criminal statutes”:

“Given all this,” Justice Gorsuch could not demand a heightened standard for criminal cases. Instead, he was convinced that “the criminal standard should be set above our precedent’s current threshold,” rather than that “the civil standard should be buried below it.”

Here would be a reason to both decide Gundy in the Petitioner’s favor and reject a special non-delegation standard for criminal cases. If this is the tack some of the Gundy justices take, we can expect to see some variant on Justice Gorsuch’s bottom line from Dimaya:

"[T]his isn’t your everyday ambiguous statute. It leaves the people to guess about what the law demands—and leaves judges to make it up. You cannot discern answers to any of the questions this law begets by resorting to the traditional canons of statutory interpretation. No amount of staring at the statute’s text, structure, or history will yield a clue. Nor does the statute call for the application of some preexisting body of law familiar to the judicial power. The statute doesn’t even ask for application of common experience. Choice, pure and raw, is required. Will, not judgment, dictates the result."

To summarize, many judges have been disturbed by the notion of legislators drafting general criminal statutes, then leaving to bureaucrats the details of deciding who goes to prison and for what. "Especially when these bureaucrats are prosecutors who are themselves “engaged in the often competitive”—and politically charged—“enterprise of ferreting out crime.” Some judges, including in the SORNA context, have proposed heightening the scrutiny demanded by the Non-Delegation Doctrine as a way of ensuring that people are only prosecuted when proper public authority demands it. Others have decided that the dire consequences to liberty and constitutional governance do not end at the threshold between criminal and civil cases, and so the

---

128 Id. at 1210 (majority opinion).
129 Id. at 1211.
130 Id. at 1212 (plurality opinion).
131 Id. at 1213.
132 Id. at 1223 (majority opinion) (citing Johnson v. United States, 135 S. Ct. 2551 (2015) (holding similar statute’s definition of “violent felony” void for vagueness)).
133 His opinion painstakingly links these two areas of law. See id. at 1227–28, Justice Thomas—who rejects vagueness as a basis for striking down laws, and so dissented in Dimaya—refused to do the same, but left the door open to non-delegation challenges over overly general statutes. Id. at 1248 (Thomas, J., dissenting).
134 Id. at 1226 (Gorsuch, J., concurring in part and concurring in the judgment).
135 Id. at 1229 (internal citation omitted).
136 Id. This could mean both standards should be raised, with the criminal one still ultimately set higher than the civil. But Justice Gorsuch focused on the similarities between many civil and criminal sanctions. Either a single standard, or different ones based on real-world consequences, seems likelier to be his preferred outcome.
137 Id. at 1232.
better route is simply to demand more exact statutes across the board. Either approach would not bode well for the government in *Gundy*.

V. Conclusion

*Gundy* may well be the case that revitalizes the Non-Delegation Doctrine. Or, it could give the Doctrine one good leg, making it very important in criminal cases but still ineffectual in virtually all civil ones. Either way, there is a good chance that it will be one of the most important criminal and administrative law cases of the early twenty-first century.
Deference to Agency Rule Interpretations: Problems of Expanding Constitutionally Questionable Authority in the Administrative State

By Ronald A. Cass

Note from the Editor:

This article argues that, while judicial deference to agency decisions is often appropriate and constitutional, Auer deference gives agencies too much unreviewable discretion such that it violates the Constitution’s separation of powers.

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.


About the Author:

Ronald A. Cass is Dean Emeritus of Boston University School of Law, Chairman and Resident Scholar at the Center for the Rule of Law, President of Cass & Associates, PC, Senior Fellow at the Center for the Study of the Administrative State, and Senior Fellow at the International Centre for Economic Research. An article exploring themes addressed in this paper (in some respects overlapping with this paper as well as giving more in-depth treatment to issues raised here) will appear in volume 87 of the Fordham Law Review and can be referenced as Ronald A. Cass, Auer Deference: Doubling Down on Delegation’s Defects Doctrine for the Modern Administrative State, 87 Fordham L. Rev. __ (forthcoming).

Together with the better-known Chevron deference rule, the doctrine articulated in *Auer v. Robbins* two decades ago—which makes reasonable administrative constructions of ambiguous administrative rules binding on courts in most circumstances—has become a focal point for concerns about the expanding administrative state. For good reason. Auer deference, even more than Chevron deference, enlarges administrative authority in ways at odds with basic constitutional structures and due process requirements.

I. Chevron: Deference from Lawful Delegation of Discretion

Chevron deference, named for the Supreme Court’s decision in *Chevron U.S.A. v. National Resources Defense Council, Inc.*, requires federal courts to defer to reasonable agency decisions implementing an agency’s statutory mandate when the particular statutory instruction being implemented is ambiguous and not clearly at odds with the agency’s actions. Although often described (and sometimes applied by reviewing courts) as if courts were directed to defer to administrators’ interpretations of law, in essence Chevron—at least as originally constructed—tells courts to decide what laws mean, and only to defer to agencies’ decisions when courts determine that Congress did not speak to an issue.

In keeping with the terms of the Administrative Procedure Act (APA), Chevron also reads the law on judicial review as directing courts to assess some administrative decisions only for their reasonableness, not their correctness. When a court concludes that Congress did not specifically instruct an administrative agency on what to do—instead granting the agency discretion with respect to some aspect of its implementation of the law—courts then should check the agency’s exercise of discretion for its reasonableness and consistency with the limits of the law, not its consistency with judges’ view of better policy. In other words, courts decide what the *law* means, including the scope of discretion granted to administrators, while administrators make *policy* decisions when Congress gave them discretion under law.

As amply documented, that was the original understanding and evident intent of the Chevron decision.

*Chevron* changed the law slightly by explicitly assuming that legislation’s ambiguity or silence on a given issue generally should be seen as granting the agency responsible for implementing that law discretion to take any action—to make any policy choice—as long as it is reasonable and not outside the scope of the law’s grant of discretion. Ambiguity, in other words, carried an implicit grant of authority for administrators to make decisions when implementing ambiguous directives—but only so far as they

---

1 519 U.S. 452 (1997).
3 While the APA provides the general rules for federal agency processes and for judicial review of agency action, the terms for review in Chevron were governed instead by the Clean Air Act. The relevant provisions in that law, however, mirror the review provisions of the APA.
did not contradict the judicially determined limits of the law. At times, the Court has said this quite clearly, as it did, for example, in *Smiley v. Citibank (South Dakota), N.A.* Even before *Chevron*, the Court at times treated legislative authorization of agency action in broad, vague terms as evidence of an understanding that Congress was giving the relevant agency (or the President) authority to make policy within the scope of those terms. *Chevron* simply generalized that inference.

Understood this way, *Chevron* allows administrators to make different policy choices over time, recognizing distinct (and constitutionally appropriate) roles for the legislature, the executive branch, and the courts. This “original *Chevron,*” thus, is consistent with the scope of authority delegated to administrators by law; it also is consistent with the Constitution’s separation of powers among the three branches.

*Chevron* should be—and certainly has been—criticized for its unclear language and its often confused and confusing application by courts (“*Chevron*-in-practice”). There is ample reason to think that *Chevron’s* direction to reviewing courts should be abandoned in favor of the terms of the APA. But original *Chevron* should be applauded, at least, for rooting its understanding of the role of administrative authority in legally prescribed delegation.

Under original *Chevron*, as under the APA, deference follows delegation. Delegation is an essential prerequisite for deference—but not necessarily the whole game.

II. *Auer* Deference: Self-Delegation and Due Process Problems

*Auer* deference as framed by the Court’s decision—what could be termed “original *Auer*”—is markedly different in kind from original *Chevron*, although a version of *Auer* deference is not (if modified to limit the doctrine’s scope, make sure it is used only when there is a statutory commitment of deference, and tailored to minimize opportunities for unfair surprise in agency interpretations). Original *Auer*—strong deference to any agency interpretation of ambiguous agency rules, no matter the nature of the ambiguity or the means or timing for a later interpretation—was represented as similar to *Chevron*. But ambiguity in rules adopted by an agency cannot plausibly be evidence of a congressional commitment of authority to the agency. And administrative officials cannot confer additional discretionary authority on themselves. If judicial deference follows from legal delegation of discretionary authority to administrators—as in original *Chevron*—that delegation must be found in statutory or constitutional provisions, not in unclear agency rules. This connection between deference and delegation is the key to understanding what is wrong with *Auer* deference and why seemingly similar deference such as original *Chevron* nonetheless is sound. Before turning further to that issue, it is worth considering other objections to *Auer* deference.

Objectors to *Auer* have given cogent reasons why courts should not grant deference to administrative interpretations merely because an agency’s rule is unclear. The most commonly voiced objections implicate, or directly invoke, due process concerns. One well-known objection—advanced notably by Professor (now Dean) John Manning—focuses on potential partiality, a corollary of permitting the body that writes rules to interpret them. The framers of the Constitution sought to avoid the potential for partiality inherent in this kind of arrangement by ensuring that legislative and judicial powers would reside with separate branches of the federal government. Another frequently voiced objection to *Auer* deference emphasizes the risk that an agency could revise its interpretation in ways that would unfairly surprise those who must comply with its rules—and might even choose less clear rule formulations for the purpose of providing leeway for different, not wholly foreseeable, applications. Judicial deference to agency rule interpretations reduces protections against these potential problems.

These objections provide a reasoned basis for skepticism about permitting administrators both to write rules and to interpret and apply rules. That skepticism, however, should not be the basis for barring any legal commitment of discretion to administrators to perform both functions and to receive deference for both. Neither objection, that is, explains why Congress should be disabled in all instances from granting administrators discretionary authority over rule interpretation, even in settings that do not carry serious risks of partiality or unfair surprise in administrative construction.

III. Administrative Discretion Explained: Legitimate Bases for Deference

There surely are legitimate reasons for granting discretion to administrators. Congress reasonably can conclude that certain decisions require confidential information, involve judgment calls on how to allocate agency resources, or will be better if informed by special expertise or experience. In such cases, delegations of decision-making authority are sensible and consistent with our constitutional structure.

Decisions necessarily based on information that cannot be widely disclosed—such as national security considerations relevant to selection, assignment, and retention of officials at the Central Intelligence Agency—properly can be assigned to the discretion of the relevant administrators. (These decisions have, in fact, been committed to CIA officials’ judgment by law.) The Supreme Court’s decision in *Webster v. Doe,* and especially the separate opinions by Justices Antonin Scalia (dissenting) and Sandra Day O’Connor (dissenting in part), recognizes the role of discretion in these judgments.

Discretionary authority also can be appropriate for determinations calling on judgments about the best use of agency resources or the best route to implement enforcement activities which require a balance of priorities and personnel and assessment of the effectiveness of alternative enforcement approaches. While prosecutorial discretion poses its own set of problems, the complex set of managerial and policy considerations relevant to prosecution decisions explains why courts, including the Supreme Court in *Heckler v. Chaney,* have deferred to administrators’ judgments on these matters.


Further, some determinations draw on technical, scientific, or experiential judgments that can be assigned to administrators’ discretion. Where in Washington, D.C., should a government building be located? Which building materials are best suited to a given structure in a particular climate? How should contracts for different quantities of and delivery schedules for concrete be compared to establish the comparability of the prices charged? What separation and siting of particular broadcast outlets best assures signal clarity and coverage? Such questions are most sensibly answered drawing on information or analysis more accessible to administrators than to reviewing judges. These sorts of questions frequently are presented in applying rules for decision as well as in framing regulations to guide future decisions.

Saying that some decisions sensibly can be committed to administrators’ discretion does not mean that any particular decision in fact has been committed to administrators’ discretion. Courts also should not assume simply because it would be reasonable to grant administrators discretion that Congress has provided statutory authority to exercise that discretion. But the fact that some decisions are sensibly committed to administrators’ discretion helps explain why sometimes deference to administrative rule implementation decisions might be required by law.

That was true in the case the Auer Court thought it was following, Bowles v. Seminole Rock & Sand Co.8 Seminole Rock was a dispute concerning application of rules for comparing prices in different contracts under the war-time Price Control Act of 1942—a dispute that should have been (and, judging from most of the language in the Court’s opinion, actually was) decided on grounds far narrower than the broad deference rule announced by the Court and later relied on as precedent for Auer. In fact, neither the Seminole Rock Court nor the Auer Court needed to invoke a rule of general deference to administrative interpretation of regulations, as both had other strong reasons for affirming the agency’s decision.

Critical readings of Seminole Rock reveal that the Price Control Act did give substantial discretionary authority to the Office of Price Administration (OPA) and that the OPA’s reading of the rule had virtually every hallmark of a rule deserving deference. It was sensible, consistent with the text and with OPA practice, announced contemporaneously with the issuance of the rule, and publicized together with the rule. In fact, Seminole Rock presented the best possible case for deference: actual delegation of discretion in OPA’s implementation authority, no significant risk of partiality in how it interpreted and implemented the agency’s rules, and no risk of unfair surprise—a risk that was uniquely absent because of the simultaneous issuance and publication of the rule and its interpretation. In other words, the case that first articulated the rule adopted uncritically in Auer elided all of the principal objections to its application.

Yet there is an additional consideration that was not discussed in either Auer or Seminole Rock: whether the delegation of authority itself is constitutional given the powers vested in each branch of government. The Constitution assigns separate powers to each of the three branches of the federal government, vesting legislative power in Congress, executive power in the President, and the judicial power in the Supreme Court and other courts created under Article III.

The terms of the vesting clauses are clear and instructive. While all three vesting clauses assign exclusive authority to one branch, unlike the unconditional assignments of executive and judicial power in Articles II and III, Article I vests in Congress only “all legislative Powers granted herein.”10 The limitation in that vesting clause emphasizes that the national government lacks the plenary powers of state governments and other inherently sovereign governments and also underscores the framers’ belief that the most fearsome and dangerous authority is the legislative power. James Madison, in Federalist No. 48, captured that sentiment in explaining the need for special constraints on the legislature, as historical observation revealed that the legislative branch was “everywhere extending the sphere of its activity and drawing all power into its impetuous vortex.”

Concern over legislative power explains why this power was subjected to special constraints. It can only be exercised through the agreement of majorities in both houses of Congress. The two houses of Congress are composed of representatives selected in different ways and at different times, to serve differently configured constituencies for different lengths of time, diffusing legislative power and guarding against ill-considered measures backed by temporary majorities moved by ephemeral passions. And the laws that survive the legislative gauntlet must be presented to the President for approval or veto. Madison, in Federalist No. 51, describes these precautions as a “remedy” for the fact that “[i]n republican government, the legislative authority necessarily predominates.”

Those who framed and ratified the Constitution took special pains to assure that the branches stayed within their assigned roles, seeking to “maintain[] in practice the necessary partition of power among the several departments.”11 They stressed that this most importantly included seeing that the carefully constructed limitations on the exercise of legislative powers would not be evaded. Congress cannot short-circuit the law-making process by lowering the vote needed to pass legislation, by allowing one house of Congress acting alone to pass laws, or by providing for law-making that bypasses presentment to the President. That is the understanding, for example, behind the Supreme Court’s decision in Immigration and Naturalization Service v. Chadha, which declared that Congress could not exercise a one-house

---

8 325 U.S. 410 (1945).
10 U.S. Const. art. I, § 1 (emphasis added).
11 Federalist No. 51 (James Madison).
veto of administrative decisions. The Court found that this veto procedure—which was defended as simply placing conditions on the exercise of lawfully granted administrative authority—amounted to making law without the constitutionally required procedures of bicameralism and presentment.

The same understanding applies to efforts to give another government official authority that is tantamount to law-making power, no matter what formal characterization is given to it. That is why the Supreme Court in Clinton v. New York found unconstitutional the Line Item Veto Act of 1996, which gave the President authority to veto three specific types of expenditure or tax benefit that had been conferred by a statute the President decided to sign into law. The Act effectively granted the President power to rewrite specific laws rather than accepting or rejecting them—and the Court appreciated that rewriting a law is no different than writing it in the first instance. The Court that decided the Clinton case understood that congressional passage of line-item veto authority was not a charitable act by which members of Congress ceded some of their power to the President. Instead, it was a means for advantaging certain interests and disadvantaged others—and, in exactly the same way, for advantaging some members of Congress and disadvantaging others. But more importantly, it was a means for evading constitutional constraints on law-making.

Together, the Chadha and Clinton cases stand for the proposition that, when Congress grants itself an exemption from ordinary law-making procedures, grants subordinate parts of the Congress law-making authority, or grants another official or entity parts of that authority, it is evading constitutionally required processes. However the evasion takes place, it is not permitted.

The lesson of Chadha and Clinton similarly undergirds the delegation doctrine (also referred to as the “non-delegation doctrine”). Although the essence of the doctrine was articulated much earlier, its classic formulation was given in Justice John Marshall Harlan’s opinion for the Court in Field v. Clark:

“That Congress cannot delegate legislative power to the President is a principle universally recognized as vital to the integrity and maintenance of the system of government ordained by the Constitution.” The separate opinion of Justice Lucius Lamar, for himself and Chief Justice Melville Fuller, similarly declared:

“That no part of this legislative power can be delegated by Congress to any other department of the government, executive or judicial, is universally recognized as a principle essential to the integrity and maintenance of the system of government ordained by the Constitution.”

Perhaps the best explanation for the delegation doctrine, however, is contained in Chief Justice John Marshall’s opinion for the Court in Wayman v. Southard: “It will not be contended that Congress can delegate to the Courts, or to any other tribunals, powers which are strictly and exclusively legislative. But Congress may certainly delegate to others, powers which the legislature may rightfully exercise itself.” Powers that fall into the first category—powers that Congress must exercise itself and cannot delegate to other officials—involve making rules on matters of such importance that they “must be entirely regulated by the legislature itself.” The second category is comprised of subjects “of less interest,” where Congress properly may make “general provisions” and leave it to others to “fill up the details.”

Unfortunately, although Wayman captures the approach taken for the nation’s first (almost) 150 years, the Court, in J.W. Hampton, Jr. & Co. v. United States, gave a different explanation of what the Constitution requires. Chief Justice (and former Chief Executive) William Howard Taft, writing for the Court, stated: “Congress has found it frequently necessary to use officers of the executive branch within defined limits, to secure the exact effect intended by its acts of legislation, by vesting discretion in such officers to make public regulations interpreting a statute and directing the details of its execution...” Distinguishing what is constitutionally permitted from what is forbidden, Taft wrote: “If Congress shall lay down by legislative act an intelligible principle to which the person or body authorized to [implement the law]... is directed to conform, such legislative action is not a forbidden delegation of legislative power.”

Following Hampton, the Court has only twice—never after 1935—found that legislation did not contain “an intelligible principle.” Even assignments that merely instructed agencies to act as “the public interest, convenience, and necessity” require or to set prices that are “generally fair and equitable” have passed the test, along with a lot of other vague, multi-faceted, amorphous directives. The result has been a virtual abandonment of serious attention to the way legislative commitments of authority to administrative officials fit (or do not fit) the Constitution’s divisions of governmental powers.

V. Auer Deference’s Larger Delegation Problem

The loss of a serious, direct judicial brake on legislative grants of power to administrators has permitted the enormous expansion of government regulation of the economy and of many aspects of health, safety, and personal behavior (retirement savings, family and child-raising decisions, and much more). Much of this regulatory structure covers subjects that long had been thought beyond the ambit of federal power.

In addition to long, detailed, and often internally inconsistent statutes, huge portions of the network of regulatory controls has come from officials who were not elected to Congress (or, indeed, elected at all) but who have been deputized to implement the laws, both through specific applications of law in particular settings and through adoption of more general, law-like rules. The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, for example, which covers more than

14 143 U.S. 649, 692 (1892) (emphasis added).
15 Field, 143 U.S. at 697 (Lamar, J., concurring in judgment) (emphasis added).
17 Id. (emphasis added).
19 Hampton, 276 U.S. at 406.
20 Id. at 409 (emphasis added).
2,000 pages, has given rise to more than 22,000 pages of implementing rules. Today's Code of Federal Regulations contains roughly 200,000 pages of agency-generated rules—approximately nine to ten times as many pages as the congressionally-passed laws collected in the United States Code—that are enforced through threat of criminal punishment, civil penalties, denial of valuable privileges, loss of benefits, and damaging publicity.

This does not mean that there is no judicial control over administrators' exercise of law-making authority. Courts have interpreted particular statutes as inconsistent with specific assertions of agency authority. The Supreme Court, for example, in Food and Drug Administration v. Brown & Williamson Tobacco Corp., rejected the FDA's assertion of power to regulate tobacco sales under the Food, Drug and Cosmetic Act. The FDA had suddenly discovered this power almost 60 years after the Act's adoption and despite the repeated failure of efforts to gain congressional approval for tobacco regulation. Courts are unlikely to defer to agency claims of authority that would effect major changes in the law without clear statutory basis, especially when the claims are inconsistent with long-standing interpretations of the law. The interpretive canon behind this inclination was pithily captured in Justice Scalia's observation that "Congress . . . does not alter the fundamental details of a regulatory scheme in vague terms or ancillary provisions—it does not, one might say, hide elephants in mouseholes."22

VI. Discretion Consistent with Constitutional Authority

Legislative commitments of power to administrators, however, include both constitutional delegations of authority—even expansive ones—and delegations that exceed the constitutional safeguards provided by bicameralism and presentment. Look first at the positive side of the ledger. Consider, for example, the authority given to the Central Intelligence Agency for certain national security matters. Among other things, the CIA's Director is instructed to "protect[] intelligence sources and methods from unauthorized disclosure."23 The Supreme Court has upheld exercises of that authority, deferring to various decisions by the Agency. It upheld the CIA's use of contracts requiring employees to protect secret information and to secure permission before publishing material that might disclose such information.24 It also deferred to the Agency's refusal of requests for release of information under the Freedom of Information Act, such as identities of sources, that the Agency deemed to be untrustworthy flows primarily from the constitutional power of the President, and Congress may surely provide that the inferior courts are not used to infringe on the President's constitutional authority.25

The Webster case involved another matter in which statutory authority grants the CIA discretion. The case addressed the CIA's decision to dismiss an employee after the Office of Security and CIA Director determined that his continued employment posed a potential threat to national security. The National Security Act of 1947 specifically grants discretion over these matters, declaring: "Notwithstanding . . . the provisions of any other law, the Director of Central Intelligence may, in his discretion, terminate the employment of any officer or employee of the Agency whenever he shall deem such termination necessary or advisable in the interests of the United States."26 The law makes clear that the determination is committed to agency discretion and that the courts should defer to the exercise of that discretion.

As Justice Scalia explained in that case, several considerations reinforce the basis for deferring by removing the question from the scope of judicial authority to review altogether. First, the issue on which review was sought was expressly committed to the administrator's discretion. Second, the law containing the directly relevant provision gave "extraordinary deference" to the Director. And, third, "the area to which the text pertains is one of predominant executive authority and of traditional judicial abstention."27 Scalia concluded that "it is difficult to conceive of a statutory scheme that more clearly reflects . . . 'commit[ment] to agency discretion by law' . . . ."28 The third consideration was central to the arguments of both Justice Scalia and Justice O'Connor in Webster. They both emphasized that it is not merely the nature of the statutory commitment but the consistency of that commitment with constitutional assignments of authority among the branches. Justice O'Connor, focusing on the concept behind the third reason for deference given by Justice Scalia, declared:

The functions performed by the Central Intelligence Agency and the Director of Central Intelligence lie at the core of "the very delicate, plenary and exclusive power of the President as the sole organ of the federal government in the field of international relations." United States v. Curtiss–Wright Export Corp., 299 U.S. 304, 320 (1936). The authority of the Director of Central Intelligence to control access to sensitive national security information by discharging employees deemed to be untrustworthy flows primarily from this constitutional power of the President, and Congress may surely provide that the inferior courts are not used to infringe on the President's constitutional authority.29

In the same vein, Justice Scalia emphasized that there are "certain issues and certain areas that [are] beyond the range of judicial review," including those so intimately bound to matters within the constitutional domain of the executive branch that insulation from review reflects "a traditional respect for the functions of the other branches."30

Determinations that are committed to agency discretion by law or are of such a nature—given constitutional assignments of executive authority and the type of discretionary judgments necessary to effectuate them—that they are incompatible with judicial review are matters on which courts should defer. This deference properly encompasses not only the act of crafting

[27] Webster, 486 U.S. at 615–16 (Scalia, J., dissenting).
[28] Id. at 616.
[29] Id. at 605–6 (O'Connor, J., concurring in part and dissenting in part).
[30] Id. at 608–09 (Scalia, J., dissenting).
regulations but also of implementing administrative authority through decisions in individual cases, through interpretation of agency rules, or through combinations of all of these activities. The sort of deference addressed in *Auer* is appropriate here—not because an agency that writes a rule should be able to interpret the rule, but because *any* agency action strictly within the executive’s domain, especially when buttressed by statutory confirmation of the understanding that this is distinctly the executive’s role, is properly within the agency’s discretion. When authority has been constitutionally delegated, courts should accept the exercise of discretion attached to that delegation and defer to it.

VII. Discretion at Odds with Constitutional Authority

In contrast, some assignments of discretionary authority to administrators plainly are in tension with—if not wholly in violation of—the constitutional division of roles for the different branches. As Chief Justice Marshall said in *Wayman* almost 200 years ago, the legislative power vested in Congress encompasses decisions on all “important subjects, which must be entirely regulated by the legislature itself . . . .” Executive branch officers may make rules for less important matters; and judges, in the course of deciding cases, may articulate rationales for their decisions that guide future decisions. But the truly important choices—those that are most politically-freighted, that are most seriously contested, and that have the most significant impact on society—cannot be left to others more than they could be made by the legislature without bicameral agreement and presentment to the President. That especially applies to choices involving “rules for the regulation of the society,” which Alexander Hamilton termed the “essence of the legislative authority.”

The point applies even more broadly. Not only is Congress charged with making big decisions about the ordering of American life, it must decide important aspects of how those big decisions will be implemented—those aspects of making law can be just as important. But the less important details of implementation surely can be assigned to others. If Congress provides benefits for veterans or for citizens with serious, work-limiting disabilities, it does not need to decide every benefit claim or prescribe every detail for who gets which services it considers to be bad and what to do about it. In the absence of a serious judicial doctrine limiting the scope of delegation, laws have effectuated large-scale transfers to administrators of authority to make rules over critically important and politically salient issues. So, for example, the Consumer Financial Protection Bureau was given authority to regulate anyone who offers or provides “a consumer financial product or service” in the broad, ambiguous terms quoted above. Worse, some agencies have asserted authority over matters doubtfully within their delegated powers. The FDA’s claimed authority over tobacco sales, rejected in the *Brown & Williamson* case, is one example. The Federal Communications Commission’s assertion of authority to regulate the pricing practices of internet service providers is another. Deferring to administrators’ decisions on such matters exacerbates the problem of judicial unwillingness to insist that important choices for regulation of private conduct be made by Congress through constitutionally-mandated processes. This point has been made by many scholars and jurists objecting to *Chevron*-in-practice: when agencies make critical policy decisions on important matters, they exercise legislative authority.

The problem is far greater, however, when deference is extended to rule interpretations in the way *Auer* requires. *Auer* expands the range of agency decisions to which courts should defer from first-level actions that directly implement constitutionally-questionable grants of authority based on statutes to second-level actions that implement agency rules. While *Chevron* provides for deference to an agency’s initial policy choices made in framing rules—primarily through “notice-and-comment” rulemaking processes designed to elicit relevant information, to allow expansive public participation in the rule-framing process, and to provide some degree of advance warning on how the agency will act—*Auer* requires deference to follow-on choices made in an array of rule interpretations and applications, generally using quite different processes that do not contain the features of notice-and-comment. If administrative law-making is problematic, allowing administrators to remake the law repeatedly—to revise the meaning of agency-made law through new interpretations of admittedly unclear agency rules—should be doubly problematic. Consider, for example, the Department of Education’s change in interpretation of a Title IX regulation respecting segregation

---


32 Federalist No. 75 (Alexander Hamilton).


34 While some older laws used similarly broad phrasing (think of the Sherman Antitrust Act of 1890, 26 Stat. 209, outlawing “every contract . . . or conspiracy in restraint of trade or commerce among the several states”), those laws tended to be (or, certainly, could and should have been) cabined by the historical usage of the terms, the terms’ common law roots, and the intentional requirements traditionally required for the courts to conclude that someone had violated such laws.

of school bathroom and locker room facilities by sex. The Department used a private letter ruling to dramatically alter an unambiguous rule some four decades after its adoption. The decision of a Fourth Circuit panel to defer to that interpretation certainly was a questionable extension of *Auer*, but it illustrates the potential range of decisions that *Auer* (or *Auer*-like) deference can entail—across time, over different vehicles for announcing changes in administrative position, and across different views of appropriate assertions of government power.

Deference to second-level agency decision-making enlarges the set of agency determinations subject to presumptive authority and vastly enlarges the set of such determinations claiming policy-based deference that are likely to be made without the sorts of procedures generally deemed best suited to informing both the administrative decision-makers and the public. The result is not only that law is made by the wrong officials without the processes constitutionally required for making law by Congress; it is made without even using the processes that agencies are supposed to utilize in writing substantive rules.

VIII. Conclusion

Examining the relationship between statutorily-directed deference and constitutional-structural principles clarifies the essential objection to *Auer* and the limits of that objection. When Congress by law confers discretionary authority that does not exceed its constitutional power to delegate functions to an administrator, courts should respect that assignment of authority unless it violates other specific constitutional commands. A different rule should apply when delegations are at most only arguably consistent with the Constitution. When that is the case, it means that the delegations probably are not consistent with the Constitution, even if they comply with the delegation doctrine as it has been interpreted by courts over the past eight decades to accommodate the modern administrative state. In this setting, deference—especially the sort of serially expanding deference *Auer* embraces to cover successive levels of administrative determination—exacerbates the problems with delegation.

A reinvigorated delegation doctrine would solve the major *Auer* problem directly, and elimination of *Auer*-like deference would be clearly preferable to retaining the doctrine in its current form. Short of that, demanding that the statutory basis for deference is clearly articulated—that Congress plainly convey authority for administrators to exercise discretion at the second level of administrative rule implementation as well as the first level of more direct statutory implementation—would provide a modest first step in cabining problems associated with constitutionally questionable delegations of law-making authority. Those who embrace the rule of law, whether advocates or opponents of the modern administrative state, should support that step.

---

Civil Rights

The Student Right to Counsel

By KC Johnson & Mike S. Adams

Note from the Editor:
This article argues that a student right to counsel in quasi-criminal campus disciplinary proceedings is good for both accused students and universities. It highlights how some states are addressing this issue through legislation, addresses opposing arguments, and surveys the status quo on campuses.

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.


About the Authors:
KC Johnson is a professor of history at Brooklyn College and the CUNY Graduate Center. Mike Adams is a professor at the University of North Carolina at Wilmington. The authors appreciate the contributions of a third unnamed author.

“He who represents himself has a fool for a client.” The advice implicit in this lawyerly adage cannot be heeded by college students involved in campus disciplinary matters addressing conduct that also could be subject to criminal penalty. Many colleges deny the right to counsel by prohibiting students’ lawyers, and sometimes the students themselves, from exercising the fundamental functions of an attorney, such as presenting evidence, cross-examining witnesses, or speaking to anyone but the client during a hearing. Such restrictions, moreover, exist at a time of unprecedented pressure—from the federal government, the media, and social activists—on colleges to adjudicate quasi-criminal behavior, especially sexual misconduct, outside the due process protections of the criminal justice system. In short, campuses have created disciplinary tribunals for quasi-criminal matters with the expectation that the accused must represent himself.

Restricting the traditional activities of legal professionals in such cases has the effect of reducing the reliability of the proceedings, which inhibits a college’s search for both justice and truth in these matters. An attorney who represents the accused can expertly interpret campus policies and hold colleges to their own stated processes throughout the proceedings. Attorneys know how to evaluate evidence and cross-examine witnesses in order to demonstrate to decisionmakers whether the evidence and the witnesses are credible. But universities do not want accused students to have robust legal representation. In 2013, defending Brown University’s policy of denying the accused student access to a lawyer in the hearing, former Brown vice president Margaret Klawunn said, “We don’t want attorneys to start running the University process.”¹

Colleges have long sought to run such proceedings under their own rules. Twenty years ago, Harvey Silverglate and Alan Charles Kors published The Shadow University: The Betrayal of Liberty on America’s Campuses, exposing the injustices that result from covert campus judicial systems that do not provide minimal safeguards of due process. In 2011, the federal government mandated further erosion of due process protections—for instance, by mandating double jeopardy should an accuser choose to appeal any case where the accused also has a right to appeal.² The new requirements codified at a national level what some colleges were already doing and what activists were seeking. Former University of Wisconsin police chief Susan Riseling, for example, suggested that administrators could use the lack of procedural protections on campus to help build criminal cases

against accused students, saying “It's Title IX, not Miranda. . . . Use what you can.”

The 2011 federal guidelines used an equity rationale based on analogy to a civil trial to require that accuser and accused be treated equally; but a disciplinary case is between the college and the accused student, more analogous to a criminal case. And while most colleges have teams of attorneys, accused students all too often are banned from having the meaningful services of an attorney during a hearing. The equity argument has not been extended—by colleges or the federal government—to a right to counsel for the accused. Although the 2011 mandates were rescinded in 2017, few colleges today voluntarily offer the right to counsel for the accused. Although the 2011 mandates were extended—by colleges or the federal government—to a right to counsel for the accused. Although the 2011 mandates were rescinded in 2017, few colleges today voluntarily offer the right to an attorney and, in fact, many college administrators actively oppose it.

The minimal role that lawyers for accused students play in Title IX tribunals has generated concern from judges at both the federal and state levels. In 2017, U.S. District Judge Philip Simon expressed puzzlement about the restrictions that Notre Dame placed on lawyers for accused students:

They can’t talk with the accused; they can’t ask questions; they can’t even pass notes to the accused. They are only permitted to consult with the students during breaks, given at the Hearing Panel’s discretion. If, for example, a witness says something very inculpatory about the accused, and there is no break, the student can’t talk with his advisor or lawyer about what he should ask the witness. And by the time the accused finally has a chance to talk with his advisor on a break, the witness could be long gone.

Similarly, in July 2018, Justice Steven Perren in California’s Second Appellate Division commented in a case against the University of California, Santa Barbara, “I read this record, and I was stunned at a university procedure which purports to be fair and equitable puts a kid [the accused student was a freshman] who’s attempting to get a college education in the position of, essentially, a lawyer in a major sexual assault case”—especially since, he noted, a representative of the university’s general counsel office participated in the hearing. This article examines the disadvantages of denying meaningful legal representation to the accused in campus misconduct cases and examines the role of state legislation in addressing these concerns.

I. The Problems with Denying Meaningful Legal Representation to Accused Students: Three Case Studies

Much has been written arguing that the post-2011 Title IX disciplinary process treats accused students unfairly. Students in many well-documented cases have been wrongly accused and found guilty in part because they lacked meaningful legal representation—or any legal representation—in the adjudication process.

Our purpose in this article is not to review this well-traveled ground, but to demonstrate how ensuring that the accused student has access to legal representation benefits both that student and the university itself—first, by making it more likely that the outcome of a disciplinary proceeding is correct, and, second, by helping to preemt expensive litigation at a later stage. As the Sixth Circuit observed:

[T]he opportunity to question a witness and observe her demeanor while being questioned can be just as important to the trier of fact as it is to the accused. “A decision relating to the misconduct of a student requires a factual determination as to whether the conduct took place or not.” . . . “The accuracy of that determination can be safeguarded by the sorts of procedural protections traditionally imposed under the Due Process Clause.”


4 Stephanie Francis Ward, Despite Title IX Guidance Repeal, Many Schools Don’t Plan to Change Handling of Sex Abuse Complaints, ABA JOURNAL (July 23, 2018), http://www.abajournal.com/news/article/despite-repeal-on-department-of-ed-title-ix-guidance-many-schools-dont-plan/ (discussing survey results indicating that only eight percent of surveyed administrators had a positive reaction to Education Secretary Betsy DeVos’ 2017 guidance, which invited schools to create fairer Title IX adjudication procedures).


7 See, e.g., KC. Johnson and Stuart Taylor, Campus Rape Frenzy: The Attack on Due Process at America’s Universities (2017); Evan Gerstmann, Campus Sexual Assault: Constitutional Rights and Fundamental Fairness (2018).

8 A 2017 survey from University of Miami Law Professor Tamara Rice Lave found that only 3 percent of the institutions she examined allowed “robust” legal representation for the accused. All other institutions that permitted an accused student to have a lawyer present during the hearing required the lawyer to remain silent. Tamara Rice Lave, A Critical Look at How Top Colleges and Universities are Adjudicating Sexual Assault 71 U. Miami L. Rev. 376 (2017), available at https://ssrn.com/abstract=2931134.

9 See, e.g., Doe v. Amherst College, 238 F. Supp. 3d 195 (2017) (college procedures denied accused student any lawyer during hearing, at which accuser admitted, without pushback from panel, to having sent texts on the night of the incident, despite having previously told the investigator otherwise); Tanyi v. Appalachian State University, 2015 U.S. Dist. LEXIS 95777 (W.D.N.C. July 22, 2015) (university assigned a lawyer to counsel the accuser during the hearing, but insisted that accused student be represented by a graduate student without any legal training); Emily Yoffe, The College Rape Overcorrection, SLATE (Dec. 7, 2014), http://www.slate.com/articles/double_x/doubles/2014/12/college Rape campus sexual assault is a serious problem, but the efforts.html (describing a University of Michigan case in which the accused student was presented with one opportunity to tell his side of the story, without legal representation and without ever appearing before a university hearing); Richard Dorment, Occidental Justice, ESQUIRE (Mar. 25, 2015), https://www.esquire.com/news-politics/a33751/occidental-justice-case/ (March 25, 2015) (describing an Occidental College case in which the accused student had to represent himself despite the police and local prosecutor having previously investigated the case and finding charges unwarranted).

10 Doe v. University of Cincinnati, 872 F.3d 393, 401 (2017) (quoting Bd. of Curators of the Univ. of Missouri v. Horowitz, 435 U.S. 78, 95 (1978)) (discussing the importance of cross-examination—to both the
The purpose of the right to assistance of counsel, like that of the right to cross-examination, is to discover the truth. Due process protections are designed to ensure, not only fair procedures, but accurate and just outcomes, which is what the university ultimately wants. Moreover, allowing accused students meaningful legal representation serves the university’s interests by providing legal input to help avoid lengthy and costly litigation down the line.

Consider, for example, a case from James Madison University. After a university disciplinary panel found an accused student not guilty in 2014, his accuser appealed to a three-person board whose members, not involved in the original hearing, reviewed the case de novo. The accused student and his legal representative had only limited rights to see the new evidence the accuser offered. The appeals board found the accused student guilty on the basis of that new evidence, and JMU suspended him for five and a half years. If university procedures had allowed a lawyer for the accused student to present evidence before the appeals board, the university’s decisionmakers would have learned that a voicemail submitted only during the appeal by the accuser, in which she sounded heavily intoxicated, actually came from a different date and thus could not have shown her intoxication on the night of the incident. Instead, it was not until her deposition in later federal litigation that one of the JMU appeals board members learned this critical piece of information. The procedural unfairness was costly: the court ultimately ordered the university to pay $849,231.25 in attorneys’ fees.

Not allowing the accused meaningful legal representation also harmed the university’s interests in a 2014 case at DePauw University. A female student claimed that she had been too intoxicated to consent to sex with a male student, Ben King, several weeks earlier. Under then-existing university policy, King could not have a lawyer accompany him even to look at the school’s investigative file, much less to speak for him at the hearing. He was found guilty. The university’s investigator focused her inquiry on determining the accuser’s level of intoxication. But DePauw defined sexual assault not as having sex with an intoxicated party, but as “engag[ing] in sexual activity with a person one knows or should know is incapacitated.” In other words, the relevant question was not whether the accused was intoxicated, but whether King knew or should have known that she was. In an order granting King’s request for a preliminary injunction, U.S. District Judge William Lawrence found “very little evidence” that King knew or should have known that the accuser—whom he had only encountered a couple of times before the incident—was incapacitated; apart from the intoxication ratings, the only evidence DePauw cited (that she was acting “giggly” and “chatty,” out of character for her) would have been meaningless to the accused student. It is possible, of course, that a student like King might detect a university misapplying its own definition of sexual assault and raise this matter during a hearing. But realistically, such attorney-like behavior is beyond the ability of most college students. If King actually had been represented by counsel, it seems likely that the lawyer would have pointed out DePauw’s errors at the beginning of the process and thus spared all parties a protracted litigation process.

A 2017 case at St. Mary’s College, a public institution in Maryland, similarly shows how allowing the accused meaningful legal representation can benefit the college as well as the accused student. Under school procedures, an investigative team produces a report into all complaints of sexual assault, which serves as the evidentiary base for an eventual decision from an outside adjudicator. When a state judge found the school’s initial adjudication violated the accused student’s due process rights, he remanded the case to St. Mary’s for another try. St. Mary’s asked its investigators to produce a new report, which went before a new adjudicator. It appeared from the adjudicator’s initial draft report as if the accused student would be found not guilty, but the adjudicator then sent a request “for any guidance you can give” to a St. Mary’s administrator, who in turn passed it on to a Maryland assistant attorney general. The resulting changes ballooned the report from four to nineteen pages, deemed two exculpatory witnesses not credible, and reversed the tentative finding of not guilty to a final finding of guilt. Under St. Mary’s procedures, the accused student’s lawyer (since the case was already in litigation, the student had hired a lawyer) could not see the adjudicator’s draft findings. If he had, he surely would have pointed out that 1) the information used to attack the exculpatory witnesses’ credibility came exclusively from the first investigation, and thus was not properly before the adjudicator, and 2) under university policies, the adjudicator was supposed to deliberate alone, rather than have assistance from a college administrator. Those facts helped ensure the accused student’s victory in court, but the lawsuit could have been avoided altogether if the student had been afforded a full right to counsel from the beginning.

It is easy to understand why universities would structure procedures to deny meaningful legal representation to accused students; doing otherwise would threaten campus administrators’ control over the adjudicatory process. But as the James Madison, DePauw, and St. Mary’s examples show, shutting out lawyers can invite litigation and frustrate the search for the truth.

II. STATE LAWS GUARANTEING A STUDENT RIGHT TO COUNSEL: THREE SO FAR

A. North Carolina: The Students & Administration Equality Act

In fall 2012, student members of the Sigma Alpha Epsilon (SAE) Fraternity at the University of North Carolina at Wilmington (UNCW) were accused of violating campus anti-hazing regulations. UNCW also charged the fraternity with providing alcohol to minors in violation of the North Carolina criminal code. Later that semester, student leaders of the fraternity were called to answer to the charges in a formal expulsion hearing to determine whether the fraternity would be derecognized.

accused and the trier of fact—in a case in which a university denied that right to the accused student).


14 Doe v. St. Mary’s College of Maryland, Civil Action No. 18-C-16-1197 (Circuit Court of St. Mary’s County, Oct. 3, 2017).
and barred from operating on campus. Although the penalty was severe, the students faced a university attorney while being denied a reciprocal right to legal counsel. Twice they requested but were denied legal representation, each time in response to questions from the administration about the alleged alcohol-related violations. Since what they said on campus could lead to misdemeanor convictions in the criminal justice system, they essentially would have to give up their due process rights to an attorney and to remain silent in order to follow UNCW’s rules.

The fraternity was suspended. Soon afterward, on February 16, 2013, fraternity chapter president Ian Gove described UNCW’s disciplinary process in a letter to all members of the North Carolina General Assembly. Gove wrote:

The administrators would not allow our student organization to be represented by a lawyer at their hearing and limited our ability to ask questions or present opposing material. Students were not allowed to have legal counsel present when university personnel brought them in for one-on-one interviews. The coercive investigative tactics used by an administrator to seek confessions caused distrust among the students and only corrupted the process. Several of the students interviewed felt intimidated or interrogated by the administrators who would repeatedly ask leading and harassing questions, some of which were considered to be inappropriate.

The standard by which UNCW administrators take disciplinary action against a student or student organization begins with a belief that you are guilty until proven innocent and requires much less evidence than “[proof] beyond a reasonable doubt.”¹⁵

At the time of Gove’s letter, the UNCW Code of Student Life permitted an accused student to have an advisor from within the university who would “advise the respondent concerning the preparation and presentation of his/her case,” but “the advisor may not be an attorney.” The advisor could “accompany the respondent to all conduct hearings.” The advisor would be permitted “access to all materials relating to the case” if the university provided the advisor, but if the student chose the advisor, the student was responsible for providing materials to the advisor.¹⁶ Most importantly, the advisor “may not be an attorney unless there are also criminal charges pending.”¹⁷ Yet since the alcohol-related charges, so far, had only been issued by the university, they were not considered to be pending criminal charges. As a result, the students were not allowed to have an attorney accompany them to disciplinary meetings or the hearing, even though the evidence produced could have been used against them in a criminal court. Gove concluded with a request that the North Carolina General Assembly pass a “Student Administrative Equality Act” that would “allow students and their organizations the right to have legal counsel present if they so choose when interacting with a university administrator [and] the right to have legal counsel represent them in any matter involving the university.”

Within weeks of receiving Gove’s letter, the state house responded. On April 10, 2013, Representatives John Bell, Rick Glazier, Nathan Baskerville, and Jonathan Jordan—two Republicans and two Democrats—filed House Bill 843. The bill was called the Students & Administration Equality Act (SAE Act). The text of the SAE Act as passed reads, in pertinent part:

§ 116-40.11. Disciplinary proceedings; right to counsel for students and organizations.

(a) Any student enrolled at a constituent institution who is accused of a violation of the disciplinary or conduct rules of the constituent institution shall have the right to be represented, at the student’s expense, by a licensed attorney or nonattorney advocate who may fully participate during any disciplinary procedure or other procedure adopted and used by the constituent institution regarding the alleged violation. However, a student shall not have the right to be represented by a licensed attorney or nonattorney advocate in either of the following circumstances:

1. If the constituent institution has implemented a “Student Honor Court” which is fully staffed by students to address such violations.
2. For any allegation of “academic dishonesty” as defined by the constituent institution.¹⁸

The law gives an officially recognized student organization the same right to have an attorney or advocate who may “fully participate” during “any . . . procedure . . . regarding the alleged violation,” with the same exception of institutions that use Student Honor Court proceedings.¹⁹

The bill passed by a 112-1 margin in the North Carolina House. On August 23, 2013, Republican Governor Pat McCrory signed the SAE Act into law.

B. Arkansas²⁰

A bill similar to North Carolina’s died in the Virginia legislature in 2014. Arkansas then became the second state to legislate the student right to counsel in April 2015 when HB 1892, originally sponsored by Representatives Grant Hodges

¹⁷ Id.
¹⁸ N.C. Code § 116-40.11 (disciplinary proceedings; right to counsel for students and organizations).
¹⁹ The Student Honor Court exception is arguably inappropriate, if not self-defeating. A student who acts as an agent of the university for disciplinary purposes is no longer acting as a student but as an administrator. Furthermore, if the Student Honor Court makes a disciplinary recommendation to the university, the process is no longer “fully staffed by students to address such violations.”
and Warwick Sabin—a Republican and a Democrat—passed with bipartisan support.

The Arkansas law provides that, except for “any allegation of academic dishonesty,” a student:

who has received a suspension of ten (10) or more days or expulsion may request a disciplinary appeal proceeding and choose to be represented at the student's expense by a licensed attorney or, if the student prefers, a non-attorney advocate who, in either case, may fully participate during the disciplinary appeal proceeding.21

In addition, if the appeal arises from a student-on-student complaint, the complaining student also may be represented in the same way.22

There are a few notable features of the Arkansas act. First, the right to an attorney only exists where the student appeals a long suspension or an expulsion, after the hearing and all witnesses and evidence have been before the adjudicator and an initial finding has been made. Second, plagiarism and other such charges of “academic dishonesty” do not produce a right to an attorney even if the penalty is expulsion. Third, the attorney may “fully participate” in the appeal, avoiding some of the problems with partial rights discussed in the case studies above. Finally, a complaining student has a right to an attorney once the accused student appeals, whether or not the accused student exercises the right to an attorney.

C. North Dakota

Within days, North Dakota became the third state to pass such a law. The bill, originally sponsored by Senators Ray Holmberg, Kelly M. Armstrong, and Jonathan Casper (as SB 2150) and by Representatives Lois Delmore, Mary C. Johnson, and Diane Larson, passed the House of Representatives 92–0 and the Senate 44–1. Governor Jack Dalrymple signed it into law on April 22, 2015. As a result, section 15-10-56 of the North Dakota Century Code now reads, in pertinent part:

Any student enrolled at an institution under the control of the state board of higher education has the right to be represented, at the student's expense, by the student's choice of either an attorney or a nonattorney advocate, who may fully participate during any disciplinary proceeding or during any other procedure adopted and used by that institution to address an alleged violation of the institution's rules or policies. This right applies to both the student who has been accused of the alleged violation and to the student who is the accuser or victim. This right only applies if the disciplinary proceeding involves a violation that could result in a suspension or expulsion from the institution. This right does not apply to matters involving academic misconduct. Before the disciplinary proceeding is scheduled, the institution shall inform the students in writing of the students' rights under this section.23

The same applies to an officially recognized student organization, and it applies to appeals under the following circumstances:

Any student who is suspended or expelled . . . and any student organization that is found to be in violation of the rules or policies of that institution must be afforded an opportunity to appeal the institution's initial decision to an institutional administrator or body that did not make the initial decision for a period of one year after receiving final notice of the institution’s decision. The right to appeal the result of the institution's disciplinary proceeding also applies to the student who is the accuser or victim. . . .

The issues that may be raised on appeal include new evidence, contradictory evidence, and evidence that the student or student organization was not afforded due process. The institutional body considering the appeal may consider police reports, transcripts, and the outcome of any civil or criminal proceeding directly related to the appeal.

Unlike the other states, North Dakota also defines “fully participate”:

“fully participate” includes the opportunity to make opening and closing statements, to examine and cross-examine witnesses, and to provide the accuser or accused with support, guidance, and advice. This section does not require an institution to use formal rules of evidence in institutional disciplinary proceedings. The institution, however, shall make good faith efforts to include relevant evidence and exclude evidence which is neither relevant or probative.24

Finally:

This section does not affect the obligation of an institution to provide equivalent rights to a student who is the accuser or victim in the disciplinary proceeding under this section, including equivalent opportunities to have others present during any institutional disciplinary proceeding, to not limit the choice of attorney or nonattorney advocate in any meeting or institutional disciplinary proceeding, and to provide simultaneous notification of the institution's procedures for the accused and the accuser or victim to appeal the result of the institutional disciplinary proceeding.

The North Dakota law also has some distinctive features that are worthy of note. First, the right to an attorney only applies to suspensions and expulsions (including of student organizations), except where those penalties are given for “academic misconduct.” Second, the right to an attorney applies both before and during appeals, unlike the Arkansas right which only applies in appeals. Third, the attorney may “fully participate . . . during any . . . procedure . . . to address an alleged violation”; “fully participate” is defined to make clear that an attorney is explicitly allowed to make opening and closing statements, examine and cross-examine witnesses, and provide other services. Also, the university must notify the accused student of the right to an attorney in writing. Finally, an accuser or victim has equivalent rights to an attorney.

22 Id.
24 Id.
III. Administrative Opposition to the SAE Act: Three Bad Arguments

In February 2015, shortly before Arkansas and North Dakota adopted their versions of the SAE Act, a coalition led by the university administrator organization NASPA: Student Affairs Administrators in Higher Education (NASPA) published an undated letter attacking the student due process movement. NASPA's letter includes some troubling statements that show why laws like the SAE Act are important countermeasures needed to combat the systemic unfairness against the accused in campus judicial procedures. Consider, for example, NASPA's argument that due process legislation would “make it more difficult for campuses to end [sexual] violence and its devastating effects on victims’ lives.” Here, the letter betrays a desire for “sending clear messages to the campus community”—a formulation from another part of the letter—through punishment. Of course, it has always been true that recognizing due process for the accused makes it harder to obtain guilty findings, while ignoring due process makes it easier to obtain findings of guilt, especially false findings. By ignoring the problem of false convictions altogether, NASPA trivializes the experiences of innocent victims of false accusations.

NASPA also argues that due process legislation “creating rights only for accused students and not student victims will enable outside interference at an unprecedented level into internal IHE [institution of higher education] administrative proceedings.” But such laws have been proposed because so many colleges have failed to provide fair proceedings, despite what the signatories call “15 years of higher education best practices.” Indeed, courts nationwide have found that colleges have violated the due process rights of accused students.26 Besides, the U.S. Department of Education’s interference in 2011, and then the Violence Against Women Act’s update with new interference from Congress in 2013 and subsequent negotiated rulemaking from the Department of Education, in addition to the Clery Act’s outside interference decades earlier, were approved warmly by the cases and policies described here.

IV. The Status Quo: Three Levels of Restrictions Colleges Place on the Right to Counsel

It is true that, under federal law, the accused and accuser each may be accompanied “by an advisor of their choice” at all disciplinary proceedings and related meetings and proceedings, in cases of alleged domestic violence, dating violence, sexual assault, or stalking.28 But the implementing regulation from the U.S. Department of Education, following a negotiated rulemaking process, expressly permits institutions to “establish restrictions regarding the extent to which the advisor may participate in the proceedings, as long as the restrictions apply equally to both parties.”29 Colleges and universities restrict the participation of attorneys so as to render their presence ineffective, as demonstrated by the cases and policies described here.

The status quo at colleges across the country limits students’ right to counsel in various ways. Some schools like the University of California, Davis flatly prohibit (or reserve the right to prohibit) the presence of legal counsel in student conduct hearings. Theses schools maintain secrecy with respect to outsiders with rules such as, “Hearings are closed except to the hearing panel or hearing officer, the accused student, the reporting party, and the witnesses … unless otherwise approved.”30 At other institutions, including the University of California, Irvine, counsel may be present at a hearing but “may be excluded from participating.”31

Other colleges, such as the College of William and Mary, only allow counsel to serve as a “silent supporter,” and then only upon two days’ prior notice. However, W&M acknowledges individual schools launched by the Obama-era Office for Civil Rights (OCR), despite the immediate, on-the-ground interference of OCR investigators.

NASPA further argues that right to counsel legislation “could perpetuate inequality between students based on who can afford an attorney,” and that “giving accused students a right to an attorney who ‘fully participates’” would upend “equality and fairness.” This statement—along with its underlying assumption that letting an accused student have a lawyer harms the accusing student—misapprehends the nature of the campus disciplinary process, where the real accuser is the college, not the student complainant (just as criminal cases are prosecuted by the government, not by victims). And the college often has a team of attorneys as well as vast resources and disciplinary experience, while the accused student is usually entering the process for the first time.

29 34 C.F.R. § 668.46(k)(2)(iv).
that the stakes are higher in a quasi-criminal case and permits participation of counsel if:

the hearing exposes him/her [the student] to potential criminal action outside the College's conduct process. The determination regarding the participation of legal counsel is final, and legal counsel will participate only to the extent authorized. Under no circumstances will the attorney be permitted to question witnesses or other parties to the proceedings, or to serve as a witness. The College may have its own legal counsel or advisor present if a student opts to have legal counsel present.32

Others, while allowing counsel to be present, bar them from participating directly in the hearing. In other words, only indirect input is permissible. For example, Pennsylvania State University and North Carolina Central University (NCCU) state that “The attorney may advise the student but may not disrupt proceedings and can be asked to leave at the discretion of the case manager [or, at NCCU, the conduct officer].”33 In student disciplinary cases at the University of California, San Diego, an attorney will be limited to communicating with their advisee and will not interrupt, disrupt, or directly participate in the Administrative Resolution meeting or Student Conduct Review, and UCSD must be given two days' notice (three days' notice for appeals) if an attorney will be present.34 In cases where a student is grieving a university action, “The grievant may be assisted by anyone, including a student advocate, but only the student advocate may speak on behalf of the grievant,” and if the student brings an attorney (five days' notice is required), "the administrative unit may be assisted by Campus Counsel or other representative it selects provided no attorney or non-student advocate shall participate directly in the proceedings.”35 At the University of California, Santa Barbara, “Students are to represent themselves. The role of the attorney or advisor is therefore limited to assistance and support of the student in making his/her own case.”36 In the California State University System, the attorney or advisor’s role is “limited to observing and consulting with, and providing support to, the Complainant or Student charged; an advisor may not speak on a Student's or Complainant's behalf.”37

The University of Florida provides that an attorney or any advisor "may not speak for the student, or address any hearing participants."38 At the University of Michigan, “Each party may be accompanied at the hearing by a personal advisor, who may be an attorney; however, the advisor may not participate directly in the proceedings, but may only advise the party. For example, the advisor may not question witnesses or make presentations.”39

And at Bucknell University, while the accused student “may be accompanied by an Adviser” of his choice, only:

the Investigator will call and question all witnesses, including the Parties. The Parties may ask the Investigator to pose additional questions or inquire further into specific matters by submitting these requests in writing or orally, at the discretion of the Chair. The Chair is empowered to reframe or disallow any questions that are irrelevant, redundant, or otherwise inappropriate.40

These restrictions ensure that the only lawyer allowed to speak in Bucknell’s hearings is the Title IX coordinator, an attorney who also serves as the university’s investigator.

These types of limitations are far from uncommon; indeed, they are the norm. The Foundation for Individual Rights in Education (FIRE) published a report last year on the state of due process on campus, analyzing the policies of the 53 top-ranked institutions (according to U.S. News and World Report).41 Of those 53 institutions, only three allow legal counsel to be present and directly participate in all non-academic disciplinary cases.42 With so few exceptions to the rule, it is no wonder that associations of student conduct administrators would oppose legislation to change this status quo.

V. Conclusion

By denying effective legal representation to accused students facing campus hearings for what would be criminal charges outside of campus, universities have created policies that not only are unfair to the accused, but that also—in the long run—can work against the university’s own interests in truth and justice in such cases. In quasi-criminal cases and other cases where significant consequences are possible (such as expulsion or

37 Timothy P. White, Executive Order 1098 (student conduct procedures), California State University (last visited July 25, 2018), https://www.csacate.edu/ro/EO-1098.html.
38 University of Florida Dean of Students Office, Advisors (last visited Apr. 17, 2018), https://sccr.dso.ufl.edu/process/advisors.
42 Those three are Cornell University, the University of Southern California, and the University of Wisconsin-Madison.
a long suspension), both accused and accuser should be allowed an attorney at all meetings, hearings, and relevant proceedings where the party is present. Full participation of attorneys—including examination and cross-examination of the parties and witnesses—will lessen the risks and costs of litigation and increase the likelihood of a process and a finding that will be accepted by all parties. If colleges do not increase due process protections for their students, it is likely that more states will follow North Carolina’s example and do so on their behalf.
The Problem with the Proliferation of Collateral Consequences

By John G. Malcolm

Note from the Editor:

This article discusses collateral consequences of criminal convictions and argues that such consequences should be rarer and more tailored to specific offenses, in order to avoid penalizing those who have already served their sentences and preventing them from reintegrating into society.

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.


I. What Are Collateral Consequences?

Today, there are more than 48,000 federal and state civil laws and regulations that restrict the activities of ex-offenders’ and curtail their liberties after they are released from confinement or their period of probation ends. Experts estimate that there are also thousands of similar restrictions in local ordinances. These restrictions are known as “collateral consequences” (as opposed

2 See id.
5 The term “ex-offender” as used in this article refers to a person with a prior criminal conviction.
6 Meek; Olivares et al.; Travis, supra note 4.
Collateral consequences are considered to be civil in nature and thus distinct from criminal laws and penalties, so courts, prosecutors, and defense attorneys have generally treated them as falling outside the scope of their control and immediate concern. Few are as aware as they should be of the full scope of these “post-sentence civil penalties, disqualifications, or disabilities” that follow a conviction, including criminal defendants and defense counsel.

Legislators have broad discretion when it comes to enacting laws creating collateral consequences. These laws are considered remedial and not punitive, and they are typically justified with appeals to public safety. They can affect, among other things, an ex-offender’s ability to get a job or a professional license; to get a driver’s license; to obtain housing, student aid, or other public benefits; to vote, hold public office or serve on a jury; to do volunteer work; and to possess a firearm.

In many cases, the public safety benefits of a particular collateral consequence significantly outweigh any burden it places on an ex-offender. For example, it is perfectly reasonable to prohibit convicted sex offenders from running day care centers or residing or loitering near elementary schools; such a prohibition is a prudent way to protect children from people with a track record of abusing the vulnerable. Prohibiting violent felons from receiving public assistance would also serve a public safety purpose. The national interest in these cases is so clear that the reasonable accommodation test of the Eighth Amendment should be applied, and the prohibition will be upheld.

Collateral consequences are considered to be civil in nature and thus distinct from criminal laws and penalties, so courts, prosecutors, and defense attorneys have generally treated them as falling outside the scope of their control and immediate concern. Few are as aware as they should be of the full scope of these “post-sentence civil penalties, disqualifications, or disabilities” that follow a conviction, including criminal defendants and defense counsel.

Legislators have broad discretion when it comes to enacting laws creating collateral consequences. These laws are considered remedial and not punitive, and they are typically justified with appeals to public safety. They can affect, among other things, an ex-offender’s ability to get a job or a professional license; to get a driver’s license; to obtain housing, student aid, or other public benefits; to vote, hold public office or serve on a jury; to do volunteer work; and to possess a firearm.

In many cases, the public safety benefits of a particular collateral consequence significantly outweigh any burden it places on an ex-offender. For example, it is perfectly reasonable to prohibit convicted sex offenders from running day care centers or residing or loitering near elementary schools; such a prohibition is a prudent way to protect children from people with a track record of abusing the vulnerable. Prohibiting violent felons from receiving public assistance would also serve a public safety purpose. The national interest in these cases is so clear that the reasonable accommodation test of the Eighth Amendment should be applied, and the prohibition will be upheld.


8 See Chin, supra note 3.

9 See, e.g., Hawker v. New York, 170 U.S. 189, 196–200 (1898); United States v. Gonzalez, 202 F.3d 20, 27 (1st Cir. 2000) (arguing that a collateral consequence, no matter how severe, is “not the sentence of the court which accept[s] the plea but of another agency over which the trial judge has no control and for which he has no responsibility.”), abrogated by Padilla v. Kentucky, 559 U.S. 356 (2010); United States v. George, 869 F.2d 333, 337 (7th Cir. 1989) (A collateral consequence "may result from a criminal prosecution, but is not a part of or entangled in the criminal proceeding.").


11 In Padilla v. Kentucky, a longtime U.S. resident and Vietnam veteran was arrested and pled guilty to transporting marijuana after defense counsel assured him that deportation would not follow a guilty plea. The federal government did institute deportation proceedings. Padilla argued he had inadequate notice of the consequences of his plea. The Supreme Court held that defense counsel must advise noncitizen defendants of potential immigration consequences of a conviction. See Gabriel J. Chin, Making Padilla Practical: Defense Counsel and Collateral Consequences at Guilty Plea, 54 How. L.J. 675 (2011); Case Comment, United States v. Muhammad: Tenth Circuit Holds that Defendant Need Not Be Informed of Collateral Consequences Before Pleading No Contest, 128 Harv. L. Rev. 1860 (2015), https://harvardlawreview.org/2015/04/united-states-v-muhammad/ (arguing that “defendants have a constitutional right to knowledge of the direct—but not collateral—consequences of their plea.”).

12 23 U.S.C. § 159 (2000) (revocation or suspension of drivers’ licenses of individuals convicted of drug offenses); see also, e.g., Fla. Stat. § 322.055(2) (same).


14 See, e.g., 20 U.S.C. § 1091(e) (prohibiting students convicted of drug offenses while receiving student aid from receiving such aid for a period of years after conviction).

15 See, e.g., 13 C.F.R. § 123.101(i) (prohibiting someone who is “presently incarcerated, or on probation or parole following conviction for a serious criminal offense,” from receiving a federal home disaster loan); 13 C.F.R. § 124.108(a)(4)(ii) (prohibiting someone who is “currently incarcerated, or on parole or probation pursuant to a pre-trial diversion or following conviction for a felony or any crime involving business integrity,” from being eligible to participate in the U.S. Small Business Administration’s 8(a) Business Development Program); see also Beitsch, supra note 13.

16 See Brian C. Kalt, The Exclusion of Felons from Jury Service, 53 Am. U. L. Rev. 65, 73–74 (2003); see also, e.g., Cal. Civ. Code § 203(a) (5) (prohibiting persons in California who have been convicted of malfeasance in office or a felony” from serving on a jury unless their rights have been restored).

17 See, e.g., Am. Bar Assoc. Comm. on Effective Crim. Sanctions & Pub. Def. Serv. D.C., Internal Exile: Collateral Consequences of Conviction in Federal Laws and Regulations 18, 31 (2009), available at https://www.americanbar.org/content/dam/aba/migrated/ccdicextralexalexthecheckdam.pdf (noting laws that bar certain offenders from volunteer work that involves the presence of a minor); Kim Ambrose, Wa. Defendant Assoc., Beyond the Conviction 12–13 (2013), http://www.defenseonet.org/resources/publications-1/beyond-the-conviction/Beyond%20the%20Conviction%20Updated%20-%202007.pdf (same); James Frank et al., Collateral Consequences of Criminal Conviction in Ohio 31, available at http://ocjs.ohio.gov/CollateralConsequences.pdf (report to the Ohio Office of Criminal Justice Services finding that Ohio law provides that “[a]ny person who has been convicted of a disqualifying offense is incompetent to hold a public office, to be publicly employed, or even to be a volunteer in certain public positions, such as volunteer firefighter.”).

purchasing or possessing firearms is another example of a targeted and tailored policy.21 Similarly, laws forcing a public official who has been convicted of bribery or public corruption to resign from office20 or prohibiting someone convicted of defrauding a federal program from participating in a related industry for a period of time impose collateral consequences that are sensible and directly related to the substance of the offenses committed.21 Other restrictions, such as those on voting, may make sense for some period of time, but perhaps not indefinitely.22 In these and other cases, the public safety justification is legitimate, and not just a cover for extending punishment of ex-offenders who have already served their sentences.

II. The Problems with Collateral Consequences

Many people convicted of crimes are never sent to prison, and, of those who are, more than 95 percent—tens of millions of people23—will eventually be released and return to our communities.24 They face long odds when it comes to trying to put their past actions behind them. In addition to having to endure the stigma associated with being convicted criminals, many ex-offenders have substance abuse issues, limited education, and limited job skills and experience. On top of these unconstitutional-congressional-overreach (“Those who are not willing to follow the law cannot claim a right to make the law for everyone else. And when an individual votes, he or she is indeed either making the law—either directly in a ballot initiative or referendum or indirectly by choosing lawmakers—or deciding who will enforce the law by choosing local prosecutors, sheriffs, and judges.”). Others, such as the NAACP, have argued that convicted felons should not lose their right to vote. See NAACP, Felon Disenfranchisement Is About Race, The Root (Oct. 2, 2012), http://www.theroot.com/articles/politics/2012/10/felon_disenfranchisement_naacp_launches_campaign/; see also Developments in the Law—One Person, No Vote: The Laws of Felon Disenfranchisement, 115 Harv. L. Rev. 1939 (2002) (criticizing felony disenfranchisement laws). State laws vary considerably on this issue, with 48 states and the District of Columbia imposing at least some restrictions on felon voting. See Nat’l Conf. of State Legislatures, Felon Voting Rights (2016), available at http://www.ncsl.org/research/elections-and-campaigns/felon-voting-rights.aspx. In Richardson v. Ramirez, 418 U.S. 24 (1974), the Supreme Court upheld the constitutionality of California’s felony disenfranchisement law. The essential issue appears to remain, as Associate Supreme Court Justice Clarence Thomas put it: Is the ex-offender “worthy of participating in civic life”? Caron v. United States, 524 U.S. 308, 318 (1998) (Thomas, J., dissenting).


22 See, e.g., 12 U.S.C. § 1829 (2000) (prohibiting persons convicted of crimes of dishonesty or breach of trust from owning, controlling, or participating in the affairs of a federally insured banking institution, subject to waiver by the FDIC; waiver may not be given for 10 years following conviction in the case of certain offenses involving the banking and financial industry; 10 U.S.C. § 2408 (2000) (persons convicted of fraud or felony arising out of defense contract prohibited from working in any capacity for a defense contractor or subcontractor for a period of at least five years); see also DiCola v. Food & Drug Admin., 77 F.3d 504, 507 (D.C. Cir. 1996) (upholding the Food and Drug Administration’s lifetime ban of a former drug company executive from “providing services in any capacity to the pharmaceutical industry” after conviction of adulterating a drug product and failing to keep adequate records: “The permanence of the debarment can be understood, without reference to punitive intent, as reflecting a congressional judgment that the integrity of the drug industry, and with it public confidence in that industry, will suffer if those who manufacture drugs use the services of someone who has committed a felony subversive of FDA regulation.”).

23 Some have argued that it is perfectly reasonable to deny the right to vote to convicted felons. See Hans A. von Spakovsky & Roger Clegg, Felon Voting and Unconstitutional Congressional Overreach, Heritage Foundation Legal Memorandum No. 145 (Feb. 11, 2015), available at http://www.heritage.org/research/reports/2015/02/felon-voting-and-
built-in challenges, they have to navigate a tangle of collateral consequences as they stake out their new lives, and the number and breadth of these consequences can be debilitating.

Regrettably, many ex-offenders will end up committing additional offenses after their release, thereby posing a continuing threat to public safety.25 Although many of these individuals would have committed additional crimes regardless of any collateral consequences imposed upon them, many others would like to turn over a new leaf and become productive, self-reliant, law-abiding members of society, but find themselves thwarted in these efforts by collateral consequences. As the American Bar Association has pointed out, “[i]f promulgated and administered indiscriminately, a regime of collateral consequences may frustrate the chance of successful re-entry into the community, and thereby encourage recidivism.”26 It is not in anyone’s best interests to consign ex-offenders to a permanent second-class status. Doing so will only lead to wasted lives, ruined families, and more crime.

A. Too Many, Too Broad, Too Opaque

Researchers for the Justice Center at the Council of State Governments have identified over 48,000 collateral consequences scattered throughout state and federal codes, with thousands more at the local level. Texas, for example, has over 200 collateral consequences in 22 different sections of the state code.27 Many other states have enacted unknown numbers of collateral consequences that are “scattered—one might say hidden—throughout their codes and regulations.”28 In addition, the number of people convicted of a crime has risen dramatically since the 1970s and, with that, the number of people living with the collateral consequences of their crimes has risen as well.

While many of the collateral consequences described above are directly targeted at promoting public safety, many others have a tenuous connection to public safety and appear to be more punitive in nature than remedial. The proliferation of such excess restrictions makes it unnecessarily difficult for ex-offenders to reintegrate into society.29 Moreover, not all collateral consequences are reasonably related to the offenses committed by those subject to them. For example, Ohio law provides for the suspension or revocation of an offender’s driver’s license upon conviction of some crimes that are entirely unrelated to driving.30 Why restrict an ex-offender’s ability to get or drive to a job or to pick up his or her children from school if that individual poses no greater risk to people on the road than any other driver? Similar problems can arise with respect to another category of collateral consequences: those that revoke eligibility for certain government benefits. For example:

- A criminal conviction may cost a military veteran his or her pension, insurance, and right to medical treatment,31 which is particularly troubling because studies indicate that veterans who are suffering from post-traumatic stress disorder and therefore in serious need of medical treatment may be more likely to commit crimes.32

- In the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, Congress barred individuals convicted of state or federal drug offenses from receiving, in addition to student aid, federal cash assistance from receiving, in addition to student aid, federal cash assistance under the Temporary Assistance for Needy Families (TANF)

25 In a study of 25,431 federal offenders released from prison or commencing a term of probation in 2005, 49.3 percent were rearrested within eight years for a new crime or for one or more technical violations of the supervised release conditions—the median time to rearrest was 21 months—31.7 percent were reconvicted, and 24.6 percent were reincarcerated. Kim Steven Hunt & Robert Dunville, U.S. Sentencing Comm., Recidivism Among Federal Offenders: A Comprehensive Overview (Mar. 2016), http://www.ussc.gov/sites/default/files/pdf/research-and-publications/research-publication/2016/recidivism_overview.pdf. In another study in 2014, 76.6 percent of offenders released from state prison were rearrested within five years, 55.6 percent were reconvicted, and 28.2 percent were reincarcerated. Matthew Durose et al., Bureau of Justice Statistics, Dep’t of Justice, Recidivism of Prisoners Released in 30 States in 2005: Patterns from 2005 to 2010 (2014), http://www.bjs.gov/content/pubs/pdf/rprts05p0510.pdf.


27 See id. at 21, 22. While some states apply collateral sanctions only to convictions rendered in that state, others apply sanctions based on convictions rendered in other jurisdictions as well, so ex-offenders must often scour the codes of multiple states if they wish to know the full scope of disabilities that might apply to them.


program and food stamps under the Supplemental Nutrition Assistance Program (SNAP).

- States may also categorically bar certain types of offenders, such as all drug and sex offenders, from government housing for any period of time, and they can suspend or revoke a driver's license on the basis of a conviction.

While these restrictions may make sense for some limited class of ex-offenders whose convictions are related to government assistance programs, depriving broad swaths of ex-offenders of the ability to get assistance for themselves and their families, to live in affordable housing in a stable environment, or to obtain educational assistance to enhance their skills is hardly conducive to helping them become productive citizens.

### B. Employment Restrictions and Recidivism

Perhaps the most ubiquitous and pernicious collateral consequences imposed on ex-offenders are restrictions on their ability to earn a livelihood. Sixty to seventy percent of the tens of thousands of identified collateral consequences are employment-related, despite the fact that employment is a top predictor of recidivism. Again, for some limited class of offenders, these restrictions may make sense. For example, federal law bars individuals with a prior criminal conviction from holding elected office, from working for the military, or in law enforcement, private security, and jobs that require a security clearance, and this limited set of restrictions makes sense insofar as it keeps those convicted of violent crimes away from weapons, and those convicted of corruption from positions of public trust.

State laws restricting employment opportunities for ex-offenders can be even more severe than federal restrictions. For example, Virginia has enacted over 140 mandatory collateral consequences that affect employment, from disqualification to hold any state "office of honor, profit, or trust" to ineligibility to hold a commission as a notary public. Ohio imposes more than 500 mandatory collateral consequences that restrict employment opportunities including employment as a contractor or truck driver.

Experts estimate that there are thousands of similar employment restrictions in local ordinances. These can bar ex-offenders from pursuing various occupations such as street peddling, cab driving, and construction. A multitude of other occupational licensing laws compounds the effect of collateral consequences insofar as they "may either explicitly exclude individuals convicted of certain criminal convictions or implicitly exclude them through a requirement that applicants be of 'good moral character,'" These include operating a dance hall, a bar, a pool hall, a bowling alley, or a movie theater, and working as a midwife, an interior designer, a barber, a contractor, an HVAC driver.

---

33 See Marc Mauer & Virginia McCalmont, Sentencing Project, A Lifetime of Punishment: The Impact of the Felony Drug Ban on Welfare Benefits (Nov. 2013, updated Sept. 2015), available at http://sentencingproject.org/wp-content/uploads/2015/12/A-Lifetime-of-Punishment.pdf. In 2015, 37 states enforced the TANF ban; 34 states enforced the SNAP ban; 25 states conditioned receipt of welfare on the nature of conviction(s) (e.g., individuals convicted of drug possession but not manufacturing or distribution may receive benefits); some looked to completion of drug treatment programs or a post-conviction waiting period. Id. at 2. See also ABA Standards, supra note 26, at 39 (arguing that prisoners themselves do not need and should not receive welfare assistance while in prison).

34 24 C.F.R. § 966.4.

35 See NACDL, supra note 30, at 33 (providing, e.g., that California bans "every person on the [sex-offender] registry" from public housing, so "those convicted of public urination in California are barred for life from public housing while those convicted of more serious violent offenses are not").


38 See Palazzolo, supra note 7.
installer or repairman, or a cab driver. The list goes on and on, even creative politicians would be hard-pressed to come up with a legitimate public safety rationale for prohibiting an ex-offender from serving as a midwife, an interior designer, an HVAC installer, or a barber. This is particularly absurd when one considers that many ex-offenders receive training to become barbers or HVAC installers and repairmen while incarcerated, only to discover upon release that they cannot get a license to practice in the one field in which they now have a marketable skill.

Research shows that states with heavy occupational licensing burdens and restrictions for ex-offenders have seen higher average levels of recidivism for new criminal offenses than have states with fewer occupational licensing burdens and restrictions. Studies have also shown a positive correlation between collateral consequences and lower employment rates as well as higher recidivism rates. Although more research is needed, existing research strongly suggests that imposing irrational restrictions on economic opportunities for ex-offenders undermines efforts to promote public safety and a cost-effective criminal justice system.

III. What Should Be Done

Like the criminal conviction itself, civil sanctions carry real consequences that can be as injurious as they are “demoralizing.” It is, therefore, time to rethink the collateral consequences we impose on people with criminal records when those consequences increase the likelihood that ex-offenders will fail in their efforts to reform and to provide for their families.

Under certain circumstances, presidents and governors can issue pardons and restore an individual’s civil rights, and courts can expunge criminal records or issue certificates of rehabilitation, thereby providing some deserving ex-offenders with some relief from the burdens otherwise imposed by collateral consequences. Employers may also help to improve ex-offenders’ employment prospects by voluntarily delaying their inquiry into a job applicant’s prior criminal record until later in the hiring process—a practice commonly referred to as a “ban the box” policy. However, it is important to this is done voluntarily—there is evidence suggesting that employers will employ race as


54 Stephen Slivinski, Turning Shackles into Bootstraps, Why Occupational Licensing Reform Is the Missing Piece of Criminal Justice Reform, Center for the Study of Economic Liberty at Arizona State University Policy Report No. 2016-01 (Nov. 7, 2016) (estimating “that between 1997 and 2007 the states with the heaviest occupational licensing burdens saw an average increase in the three-year, new-crime recidivism rate of over 9%. Conversely, the states that had the lowest burdens and no [good-character] provisions saw an average decline in that recidivism rate of nearly 25.%.”).

55 See Sohoni: Ugen & Manza; Seiter & Kadel, supra note 29.


a proxy for criminality when “ban the box” is mandated by the government.59

There are also things state and federal legislators can do to address unduly onerous collateral consequences. Legislators should consolidate all existing collateral consequences in a single location in order to make them more accessible so the public (including defendants and their attorneys) is aware of the full consequences of criminal conviction.60 In addition, legislators should reassess the collateral consequences enacted within their jurisdictions to ensure that they are necessary to protect the public, reasonably related to the offense committed, and not capable of being enforced indiscriminately or arbitrarily. Any restriction that does not satisfy these parameters should be amended or repealed so that ex-offenders who are earnestly working to lead lawful, prosperous lives and to provide for their families are not needlessly thrown off-course.61 Legislators might also consider establishing more robust procedures for ex-offenders to petition for relief or waivers from certain collateral consequences, which could be granted in meritorious cases.

IV. Conclusion

In light of growing evidence that a number of collateral consequences may frustrate reintegration into the community and encourage recidivism, some states have already begun to reassess what collateral consequences should attach to which convictions, as well as why and for how long.62 While some collateral consequences are justifiable as a way to protect public safety, many are not. Unjustifiable collateral consequences are punitive in nature, designed to continue punishing ex-offenders once they complete their sentences for the crimes they committed. The public’s desire to continue to stigmatize an ex-offender may be understandable, but it comes at a high cost and should be resisted to promote justice and public safety.

Since most ex-offenders—millions of them—at some point will be released from custody and return to our communities, it is important that we do everything we can to encourage them to become productive, law-abiding members of society and that we not put too many impediments, in the form of excessive collateral consequences, in their way that will hinder their efforts. More attention must be paid to this issue to avoid these dangerous and counterproductive results. In a time of intense polarization, this is one of the few issues people can rally around and on which we can find common ground. It is not in anybody’s best interest to relegate the formally incarcerated to a backwater of second-class citizenship status.


60 See ABA Standards, supra note 26 (standard 19-2.1); see also Margaret Colgate Love, Collateral Consequences After Padilla v. Kentucky: Punishment to Regulation, 31 St. Louis U. Pub. L. Rev. 87, 118–21 (2012) (arguing that counsel should inform defendants of potential collateral consequences).

61 See generally ABA Standards, supra note 26; see also Dep’t of Justice, Smart on Crime—Reforming the Criminal Justice System for the 21st Century 5 (Aug. 2013), https://www.justice.gov/sites/default/files/ag/legacy/2013/08/12/smart-on-crime.pdf. Some organizations, such as the National Association of Criminal Defense Lawyers, have suggested an even more aggressive approach to addressing the problems created by overweening collateral consequences. See NACDL, supra note 30, at 33.

62 See Vera, supra note 10 (on state reform efforts between 2009-2014).
This short article examines how the extraterritoriality doctrine might apply to state energy taxes imposed on electricity that is generated in one state but used in a different taxing state, when the purpose of the tax is to discourage greenhouse gas emissions. The extraterritoriality doctrine precludes a state from regulating commerce occurring wholly outside the state's borders.1 This limitation on states’ ability to regulate commerce beyond state lines stems from the Supreme Court's dormant commerce clause jurisprudence, which establishes that the Commerce Clause implicitly curtails state regulation of interstate commerce.2

My objective is to highlight some recent cases applying the extraterritoriality doctrine in order to explore how a court might analyze the constitutionality of a state “carbon” tax on imported electricity, recognizing that while state laws may terminate at state lines, the electrical grid does not. I specifically focus on Washington state because it has recently been a hotbed of activity on energy taxes, including several proposals for carbon taxes that would have taxed the sale of coal-based electricity generated at power plants located beyond Washington’s borders. Moreover, I hope to show that while the extraterritoriality doctrine itself may not be uniformly embraced in the courts, case law counsels that the doctrine should be respected when evaluating the legality of state taxes on imported electricity.

Part I covers two recent Court of Appeals cases that stake out markedly different approaches to the extraterritoriality doctrine. With these cases in the background, Part II looks at proposed Washington state taxes on imported electricity that might become a blueprint for similar efforts in other states. Part III discusses the Ninth Circuit’s precedent on the extraterritoriality doctrine, finding that the doctrine is very much alive in that circuit—the circuit where a challenge to a state energy tax is most likely to occur. Part IV concludes with some remarks about the way a constitutional challenge to a state tax on imported electricity might unfold.

I. The Extraterritoriality Doctrine: Dead or Alive?

A. Energy and Environment Legal Institute v. Epel

In the 2015 case Energy and Environment Legal Institute v. Epel, then-Judge Neil Gorsuch wrote an opinion for the Tenth Circuit upholding Colorado’s renewable energy law against a dormant commerce clause challenge.3 The state law at the center of Epel required electricity suppliers in Colorado to ensure that a portion of the electricity they sold to Coloradans was generated with renewable resources.4 The plaintiffs challenging the Colorado...
law brought suit under the extraterritoriality doctrine of the
dormant commerce clause.5

The Tenth Circuit identified a series of U.S. Supreme
Court opinions deeming “almost per se invalid” a category
of state laws that control conduct taking place beyond the
geographic boundaries of the state.6 These opinions, which form
the foundation of the extraterritoriality doctrine, establish a
constitutional test that has been framed in various ways, asking
whether a state law has “the practical effect of . . . control[ling]
conduct beyond the boundary of the state,”7 whether a state is
“project[ing] its legislation”8 into another state, or whether a state
law regulates prices in out-of-state transactions.9

While acknowledging the viability of the extraterritoriality
document, the Epel court expressed remarkable skepticism about
it, opining that the doctrine is the “least understood” and “most
dormant” strand of dormant commerce clause jurisprudence.10
Indeed, the extraterritoriality doctrine did not help the plaintiffs in
Epel, as the court only grudgingly conceded that the doctrine
might still exist.11 The court limited the extraterritoriality doctrine
to cases involving price controls—which did not undermine
Colorado’s energy law—and closed the book on the lawsuit.12

B. North Dakota v. Heydinger

Reading Epel in isolation would create the impression that the
extraterritoriality doctrine is on its way to obsolescence. But
another opinion involving a state energy law, published a year after
Epel and in a different circuit, rejuvenated the extraterritoriality
document.13

In North Dakota v. Heydinger, Judge James Loken penned
the lead opinion invalidating a Minnesota greenhouse gas
emissions statute.14 Several electricity suppliers who wanted to
do business in the Midcontinent Independent System Operator (MISO)
territory—a regional electric grid that includes Minnesota—struggled to arrange power purchase agreements to
supply electricity generated by coal-fired plants located outside
Minnesota due to concerns that such agreements would violate
a Minnesota statute prohibiting the importation of electricity
into Minnesota from certain facilities that contribute to carbon
dioxide emissions.15 The statute in Heydinger did not survive under
the extraterritoriality doctrine, according to Judge Loken.16 He
applied a rule declaring that a state statute is invalid per se if its
practical effect is to control conduct beyond the boundaries of the
state, which includes requiring people or businesses to conduct
their out-of-state commerce in a certain way.17

As the Eighth Circuit saw it, the practical effect of the
Minnesota statute was to reduce emissions occurring outside
Minnesota by prohibiting transactions that originated at a
generation source outside Minnesota.18 This extraterritorial
control, said Judge Loken, meant that other states in the MISO
region, which are interconnected by an electric grid transmitting
electricity from numerous generation sources in a manner that is
impossible to trace, would have to abide by Minnesota’s policy
whenever generating capacity was added to the regional grid.19
According to Judge Loken, “[t]his Minnesota may not do without
the approval of Congress.”20

II. State Energy Taxes

Washington state has been center stage in a public policy
debate about proposals to impose state taxes on greenhouse gas
emissions. Numerous bills seeking to establish the nation’s first
carbon tax in Washington have been introduced in the legislature
in recent sessions,21 and a 2016 ballot initiative failed to pass.22

5 Epel, 793 F.3d at 1172. The Epel court traced the origin of the
extraterritoriality doctrine to Baldwin v. G.A.F. Seelig, Inc., 294 U.S. 511,
521 (1935). Epel, 793 F.3d at 1172. See also Chad DeVaux, One T oke
Too Far: The Demise of the Dormant Commerce Clause’s Extraterritoriality
Doctrine Threatens the Marijuana-Legalization Experiment, 58 B.C. L.
Rev. 953, 962-67 (2017) (tracing development of the extraterritoriality
document in case law); Tessa Gellerson, Extraterritoriality and the Electric
Grid, North Dakota v. Heydinger, A Case Study for State Energy
Regulation, 41 Harv. Envtl. L. Rev. 563, 569-81 (2017) (same); David
M. Driesen, Must the State Discriminate Against Their Own Producers
Under the Dormant Commerce Clause?, 54 Hous. L. Rev. 1, 15-30
(2016) (same).

6 Baldwin, 294 U.S. at 521.

(citing Healy, 491 U.S. at 324). Accord Ass’n for Accessible Medicines v.
Frosh, 887 F.3d 666, 667-74 (4th Cir. 2018).

8 Epel, 793 F.3d at 1172. Further, the Epel Court questioned whether the
extraterritoriality doctrine really established a separate test under the
dormant commerce clause. Id. at 1173. Accord New York Pet Welfare
Ass’n v. City of New York, 850 F.3d 79, 91-92 (2d Cir. 2017).

9 Epel, 793 F.3d at 1173.

10 Id.

11 North Dakota v. Heydinger, 825 F.3d 912, 913-23 (8th Cir. 2016)
(Loken, J.).

12 Heydinger, 825 F.3d at 913-14.

13 Id. at 916-17. The MISO controls over 49,000 miles of transmission lines
on a grid that covers fifteen states and parts of Canada. Id. at 915.

14 Id. at 919-22.

15 Id. at 919 (quoting Cotto Waxo Co. v. Williams, 46 F.3d 790, 793 (8th
Cir. 1995)). Additionally, Judge Loken disagreed with Epel’s conclusion
that the extraterritoriality doctrine only applies to price control statutes.
Heydinger, 825 F.3d at 920.

16 Id. at 921.

17 Id. Judge Loken found that generators in MISO cannot prevent energy
they place on the grid to serve non-Minnesota customers from being
imported into Minnesota, and a Minnesota electricity supplier cannot
do business with out-of-state generators without importing electricity
from their coal-fired facilities. Id. But cf. Elissa Walter, Flaw or Oscillate?:
The Mismatch Between the Language Judges and Attorneys Use to
Describe Electricity and the Actual Behavior of Electricity on the Grid, 44
Ecology L.Q. 343, 362-65 (2017) (arguing that Judge Loken’s analysis
misunderstands the electric grid).

18 Heydinger, 825 F.3d at 922.

19 Id. Judge Loken found that generators in MISO cannot prevent energy
they place on the grid to serve non-Minnesota customers from being
imported into Minnesota, and a Minnesota electricity supplier cannot
do business with out-of-state generators without importing electricity
from their coal-fired facilities. Id. But cf. Elissa Walter, Flaw or Oscillate?:
The Mismatch Between the Language Judges and Attorneys Use to
Describe Electricity and the Actual Behavior of Electricity on the Grid, 44
Ecology L.Q. 343, 362-65 (2017) (arguing that Judge Loken’s analysis
misunderstands the electric grid).

20 Heydinger, 825 F.3d at 922.

21 E.g., SB 5385 (2018); SB 5509 (2018); SB 6096 (2018); SB 6203 (2018);

22 Secretary of State Kim Wyman, November 8, 2016 General Election
Results, Initiative Measure No. 732 Concerns Taxes, (April 10, 2018, 9:14 AM),
Some of these proposals would have imposed a tax on emissions attributable to imported electricity, meaning electricity that is generated outside the state of Washington.\textsuperscript{23} Imposing a tax on emissions attributable to imported electricity requires the taxing state to calculate and assign a value representing the emissions created by the production of electricity at an out-of-state generation source.\textsuperscript{24} The tax would be collected from a legally responsible party at the first taxable transaction in the state, such as a utility supplying electricity in Washington, but the real targets of such taxes—the entities whose behavior they are meant to affect—are the out-of-state generation facilities that burn fossil fuels.\textsuperscript{25}

Energy tax proponents hope that such taxes will discourage the use of heavily taxed products because those products, such as coal-based electricity, have greater emissions.\textsuperscript{26} However, a grand design to transform a regional energy market by shaping behavior through taxation at the state level raises legal questions about any single state’s ability to project its policy preferences into other states.\textsuperscript{27} This is especially important for Washington to consider: Washington is home to only one coal-fired power plant, which is scheduled to shut down.\textsuperscript{28} But some of Washington’s neighboring states in the West have more coal plants and continue to serve Washington customers.\textsuperscript{29} An energy tax that speeds the demise of another state’s coal facilities could prompt close examination of controlling commercial conduct taking place in other states. A state tax that is intended to remake a regional energy market—particularly a tax that purposefully increases the price of a specific product originating in another state in order to drive that product into extinction—will be scrutinized under the extraterritoriality doctrine in courts that adopt the \textit{Heydinger} approach.\textsuperscript{30}

\section*{III. The Extraterritoriality Doctrine in the Ninth Circuit}

\subsection*{A. Rocky Mountain Farmers Union v. Corey}

In addition to reviewing Justice Gorsuch’s narrow approach in the Tenth Circuit and Judge Loken’s broader approach in the Eighth Circuit, anyone wanting to know how the extraterritoriality doctrine might apply in Washington or other west coast states should look to Ninth Circuit opinions.

The Ninth Circuit’s 2013 opinion in \textit{Rocky Mountain Farmers Union v. Corey} is a good place to start.\textsuperscript{31} \textit{Corey} was a dormant commerce clause challenge to California’s low carbon fuel standard program.\textsuperscript{32} Under that program, California regulated the carbon intensity of various fuels used for transportation, requiring fuel producers to meet state benchmarks for greenhouse gas emissions attributable to fuel.\textsuperscript{33} Determining the carbon intensity of a given fuel involved a lifecycle analysis of the emissions associated with that fuel based on numerous factors, including the efficiency of production, source of electricity used at the production facility, and whether land was converted for production.\textsuperscript{34} Many of these activities associated with fuel production took place before the fuel entered California.

Among other claims, the plaintiffs in \textit{Corey} contended that the low carbon fuel standard impermissibly regulated extraterritorial conduct in violation of the dormant commerce clause.\textsuperscript{35} In analyzing that claim, the court confirmed that the extraterritoriality doctrine is a cognizable legal theory in the Ninth Circuit.\textsuperscript{36} The court tethered its opinion to the Supreme Court’s \textit{Healy} decision, which established that the critical inquiry for extraterritoriality analysis is “whether the practical effect of the regulation is to control conduct beyond the boundary of the state.”\textsuperscript{37} The Ninth Circuit also relied on the Supreme Court’s decision in \textit{C & A Carbone, Inc. v. Town of Clarkstown}—a case involving an ordinance that required waste to be processed at a town’s transfer station—as an example showing that

\begin{itemize}
\item 23 \textit{Heydinger}, 825 F.3d at 913-23.
\item 24 \textit{Id.}
\item 25 \textit{Id.}
\item 26 Jay Inslee, \textit{Our State, Our Destiny 6-7} (Jan. 9, 2018) (“It is time to step up and give our citizens what they demand and deserve . . . which is a fight against climate change and the damaging health effects of carbon pollution . . . Now is the time to join in action and put a price on carbon pollution.”).
\item 27 \textit{Heydinger}, 825 F.3d at 913-23.
\item 28 \textit{Id.}
\item 29 \textit{Id.}
\item 30 \textit{Id.}
\item 31 \textit{Epel}, 793 F.3d at 1172-73.
\item 32 \textit{Heydinger}, 825 F.3d at 913-23.
\item 33 \textit{Id.}
\item 34 \textit{Id.}
\item 35 \textit{Id.}
\item 36 \textit{Id.}
\item 37 \textit{Id.}
\item 38 \textit{Id.}
\item 39 \textit{Id.}
\item 40 \textit{Id.} at 1101 (quoting \textit{Healy}, 491 U.S. at 336).
\end{itemize}
the extraterritoriality doctrine may apply to cases involving environmental regulations and is not limited to price control laws. And the Ninth Circuit even endorsed the plaintiffs’ assertion that a state’s police power does not allow it to “invade [another state] to force reductions in greenhouse gas emissions.”

But that is as far as the Corey plaintiffs could carry their case. Applying the extraterritoriality doctrine, the Ninth Circuit found that California’s low carbon fuel standard did not have the practical effect of controlling conduct outside the state. The court instead viewed the California program as a system that probably will influence out-of-state fuel producers as they make commercial decisions about their fuel blends, but which does not actually mandate compliance with any particular California policy in out-of-state transactions.

B. Sam Francis Foundation v. Christie’s, Inc.

Nevertheless, the Corey plaintiffs’ failure to prevail under the extraterritoriality doctrine should not be taken to mean that the Ninth Circuit will always reject claims pursued under the doctrine. Two years after Corey, the Ninth Circuit, sitting en banc in Sam Francis Foundation v. Christie’s, Inc., fully embraced the extraterritoriality doctrine and used it to partially strike down a California law.

The ill-fated state law in Christie’s was the California Resale Royalty Act. The Act required a seller of fine art to pay five percent of the sale price to the artist if the seller resided in California. The Ninth Circuit “easily” concluded that the Act violated the dormant commerce clause under the “simple” and “well established” extraterritoriality doctrine. As the court explained, Supreme Court precedent provides that the dormant commerce clause precludes state statutes that regulate commerce beyond a state’s borders. The California law in Christie’s squarely fell within the extraterritoriality doctrine’s perimeter because it directly regulated some art sales that would occur entirely outside California.

Christie’s stands as strong confirmation that the Ninth Circuit will apply the extraterritoriality doctrine to test state laws’ compliance with the dormant commerce clause. Moreover, the Ninth Circuit’s opinions in Corey and Christie’s show that it will apply the extraterritoriality doctrine in a wide variety of cases, in contrast to the Tenth Circuit’s tightly circumscribed approach limiting the doctrine to price controls.

IV. Conclusion

This brief survey of recent cases applying the extraterritoriality doctrine highlights one important constitutional consideration relating to state energy taxes on imported electricity. Taxing coal-based electricity in order to make it less competitive might look to some like sound environmental policy, but it can also be seen as one state’s unconstitutional push to regulate interstate commerce by asserting control over out-of-state facilities. A state tax on imported electricity may be especially vulnerable to this critique in a state like Washington that has comparatively few coal-fired facilities vis-à-vis its neighbors—even more so if the patent purpose of the tax is to wear down a particular industry that mainly operates beyond Washington’s borders.

There are counterarguments. For one, the Epel decision may have weakened the extraterritoriality doctrine as applied to energy laws. Another court might adopt Epel’s reasoning and conclude that a state energy tax on imported electricity—like a renewables mandate—does not raise constitutional red flags. That scenario, however, is unlikely to play out in the Ninth Circuit, where Corey and Christie’s demonstrate that the extraterritoriality doctrine still holds sway.

Alternatively, a court might determine that a state energy tax on imported electricity is analogous to the program upheld by the Ninth Circuit in Corey—merely a decision by one state to pay a price for the “ill effects” of its electricity consumption, which incidentally affects interstate commerce without unconstitutionally regulating it. This outcome is difficult to predict because it would largely depend on the details of a specific case and how the challenged state law actually functioned in relationship to other states.

In the end, if a case is eventually presented, a court will be asked to decide whether a state tax on imported electricity that seeks to phase out a specific generation resource primarily used out of state violates the dormant commerce clause. The answer to that question will have major ramifications for state regulatory authority in the energy arena going forward.

---

41 Id. at 1102 (quoting C & A Carbone, Inc. v. Town of Clarkstown, 511 U.S. 383, 393 (1994)). “States and localities may not attach restrictions to exports and imports in order to control commerce in other States.”

42 Corey, 730 F.3d at 1103 (quoting Massachusetts v. EPA, 549 U.S. 497, 519 (2007)).

43 Id. at 1106.

44 Id.

45 784 F.3d 1320, 1322 (9th Cir. 2015).

46 Id.

47 Id.

48 Id. at 1323, 1325.

49 Id. at 1323-25.

50 Id. To explain the workings of the law, the court hypothesized that the California law would require a California resident temporarily living in New York who purchased art from a North Dakota artist in New York and then sold the art to her friend in New York to remit five percent of the sale price to the North Dakota artist. Id. at 1323.

51 Epel, 793 F.3d at 1173.

52 See Heydinger, 825 F.3d at 919-22.

53 Epel, 793 F.3d at 1173.

54 Christie’s, Inc., 784 F.3d at 1323-25; Corey, 730 F.3d at 1101-06.

55 Corey, 730 F.3d at 1106.
A Shy Frog, the Administrative State, and Judicial Review of Agency Decision-Making: A Preview of Weyerhaeuser v. United States Fish & Wildlife Service

By Mark Miller

Note from the Editor:
This article previews one of the first cases of the Supreme Court’s upcoming term: Weyerhaeuser v. U.S. Fish & Wildlife Service. The article summarizes the parties’ positions and indicates a preference for a ruling in favor of the landowners challenging the Service’s designation of an unoccupied critical habitat for the endangered dusky gopher frog.

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.


II. Background on the Endangered Species Act and Critical Habitat Designations Under the Law

Congress passed the Endangered Species Act (ESA) in 1973. It recognized that “various species of fish, wildlife, and plants in the United States ha[d] been rendered extinct as a consequence of economic growth and development untempered by adequate concern and conservation,” and it thus pledged—through the application of the ESA—to “conserve to the extent practicable the various species of fish or wildlife and plants facing

3 Markle Interests, LLC v. United States Fish & Wildlife Service, 827 F.3d 744, 761 (D. Ct. E.D. La. 2014) (“Indeed it [the frog] hasn’t been sighted there since the 1960s.”).
5 Markle Interests, LLC, 40 F. Supp. 3d at 763 n.29 (“the last observation of a dusky gopher frog in Louisiana was in 1965”).
6 Markle Interests, LLC v. United States Fish & Wildlife Service, 40 F. Supp. 3d 744, 761 (D. Ct. E.D. La. 2016) (“Indeed it [the frog] hasn’t been sighted there since the 1960s.”).
7 The Landowners did not challenge the critical habitat designation as it relates to the Mississippi properties so designated. Markle Interests, LLC, 827 F. 3d at 459.
Extinction.11 Section 4 of the ESA requires the Secretary of the Interior (Secretary) to list a species as “endangered” when it “is in danger of extinction throughout all or a significant portion of its range.”12 Section 9 prohibits any person from harassing, harming, or capturing an endangered species, and it may prohibit habitat modification.13

Under Section 4 of the Endangered Species Act, when a species is listed as threatened or endangered, the Service must designate critical habitat for that species “to the maximum extent prudent and determinable.”14 The designation must be based on “the best scientific data available” and may only be made after the Secretary considers and weighs the cost of all relevant impacts, including economic impacts.15 In 1978, Congress amended the ESA to define “critical habitat”:

(i) the specific areas within the geographical area occupied by the species, at the time it is listed in accordance with the provisions of section 1533 of this title, on which are found those physical or biological features (I) essential to the conservation of the species and (II) which may require special management considerations or protection; and

(ii) specific areas outside the geographical area occupied by the species at the time it is listed in accordance with the provisions of section 1533 of this title, upon a determination by the Secretary that such areas are essential for the conservation of the species.16

Subsection (i) defines critical habitat in terms of the physical and biological features the area must possess.17 Subsection (ii) provides for the designation of unoccupied critical habitat, but only where the Secretary determines that the area is “essential for the conservation of the species.”18 Since the Louisiana property is unoccupied by the frog, both of these subsections are at issue in the case.

II. Conserving the Dusky Gopher Frog on Non-Habitat Land Would Be Expensive

In designating critical habitat for the frog in Louisiana and Mississippi, the Service identified three “primary constituent elements” (PCEs), which are defined by regulation as “the principal biological or physical constituent elements [within a defined area] that are essential to the conservation of the species.”19 These three PCEs include: (1) “small, isolated, ephemeral, acidic breeding ponds having an open canopy,” (2) upland forests “historically dominated by longleaf pine, adjacent to and accessible to and from breeding ponds, that are maintained by fires frequent enough to support an open canopy,” and (3) “[a]ccessible upland habitat.”20

The land in Mississippi designated critical habitat contains those three essential characteristics; the Louisiana land does not—it contains, at most, only the ephemeral pond characteristic described in the first PCE.21 Nevertheless, the Service defended its decision to designate the Louisiana property by asserting that, in the event of a catastrophic event in Mississippi, the Louisiana property could serve as habitat for the frog, with significant changes to create the other two PCEs.22

Pursuant to the ESA, the Service must “take[e] into consideration the economic impact . . . of specifying any particular area as critical habitat,” and it “may exclude any area from critical habitat” based on economic impacts.23 Before the final rule designating the Louisiana land was published, the Service prepared a final Economic Analysis24 analyzing the potential economic impacts associated with the designation of critical habitat for the dusky gopher frog.25 The Economic Analysis considered three possible scenarios and ultimately concluded that the designation of the Louisiana property alone could result in lost development value of $33.9 million.26 Meanwhile, the impact on the Mississippi critical habitat designations would amount to, at most, $102,000.27 This lopsided economic impact resulted from the fact that the Mississippi critical habitat is already actively managed for the recovery of the frog, while the Louisiana property is not.28

Despite the drastic economic impact and the lack of biological benefit to a frog that could not survive on the Louisiana land, the Service designated it critical habitat. That designation prompted the Landowners’ lawsuits that led to the current Supreme Court case.

III. Procedural History of the Case

The Landowners filed separate lawsuits and sought identical declaratory and injunctive relief.29 They alleged the rule designating their Louisiana property (not the Mississippi property) violated the ESA and the Administrative Procedure Act (APA), the Commerce Clause of the federal Constitution,

19 Id. at 762 (citing 50 C.F.R. § 424.12(b) (emphasis added)).
21 Markle Interests, LLC, 40 F. Supp. 3d at 761.
29 Markle Interests, LLC, 40 F. Supp. 3d at 748.
and the National Environmental Procedure Act (NEPA). 30 The Center for Biological Diversity and the Gulf Restoration Network were granted leave to intervene as defendants. 31

Upon cross-motions for summary judgment, the district court held that the Service had acted within the law in designating the Louisiana property critical habitat. 32 But Judge Martin L. C. Feldman did not mince words in describing his view of the Service’s designation of the Louisiana property, calling the Service’s actions “odd,” “troubling,” and “harsh,” and remarking that “what the government has done is remarkably intrusive and all the hallmarks of governmental insensitivity to private property.” 33 Nevertheless, considering himself to be “restrained” by the “confining” and “somewhat paralyzing” standard of review under the APA, Judge Feldman reluctantly affirmed the critical habitat designation as within the delegated powers of the agency pursuant to the ESA. 34 The district court also rejected the Commerce Clause challenge and other arguments made by the Landowners. 35

The Fifth Circuit affirmed in a 2-1 split opinion. 36 The panel majority concluded that the Service’s designation of the Louisiana property was entitled to Chevron deference because Congress did not define “essential” habitat as it concerns unoccupied critical habitat and thus delegated the definition to the Service. 37 The majority also rejected the argument that the Service should have excluded the Louisiana property because of the disproportionate economic impacts the Landowners would suffer from its designation, concluding that the Service’s decision on that point was wholly discretionary and unreviewable. 38 The Court also rejected the other arguments made by the Landowners. 39 In her dissent, Judge Priscilla Owen observed that the designated area is not essential for the conservation of the species “because it plays no part in the conservation” of the species. 40 As she put it, “[t]here is no evidence of a reasonable probability (or any probability for that matter)” that the designated area will ever become essential to the conservation of the species. 41

The full court rejected the Landowners’ motion for en banc review with an 8-6 vote. 42 Writing for the six-member dissent, Judge Edith Jones argued that the Service’s actions in this case fell far outside the authorization of the ESA: “The panel opinion...approved an unauthorized extension of ESA restrictions to a 1,500-acre-plus Louisiana land tract that is neither occupied by nor suitable for occupation by nor connected in any way to the [dusky gopher frog].” 43 The dissent was troubled by the fact that “[n]o conservation benefits accrue to [the frog], but this designation costs the Louisiana landowners $34 million in future development.” 44 From the panel decision and the denial of en banc review, the Landowners sought review.

IV. The Questions Before the Supreme Court

The Supreme Court granted review 45 to consider two questions: (1) whether the ESA prohibits designation of private land as unoccupied critical habitat if it is neither habitat nor essential to species conservation, and (2) whether an agency decision not to exclude an area from critical habitat because of the economic impact of designation is subject to judicial review. 46 These two questions are fundamentally about how far an agency like the Service can reach in filling in the gaps in statutes written by Congress, and whether this agency decision-making is insulated from judicial review.

A. What Does “Essential” Mean?

The Service and the Landowners disagree about the scope of the authority the ESA gives the Service to protect an endangered species. How much private property can the Service cordon off from private use in the name of meeting the goals of the ESA? The arguments on both sides demand careful consideration from anyone who takes both the ESA and government power seriously.

1. The Service’s Argument: Congress Asks the Service to Protect Endangered Species, and This Critical Habitat Designation Protects the Endangered Dusky Gopher Frog

In order to accomplish the underlying goal of the ESA—the conservation of endangered species—the lower courts and the Service relied upon the wide latitude the APA and Chevron deference give the Service in carrying out its statutory mission. Their arguments flow from the general proposition that the Service

30 Id. at 752-53.
31 Id. at 753.
32 Id. at 769.
33 Id. at 759.
34 Id.
35 Id. at 765.
36 Id. at 759.
37 Id. at 759-60.
38 Id. at 760-69.
39 Markle Interests, LLC, 827 F.3d 452.
41 Markle Interests, LLC, 827 F.3d at 467-72.
42 Id. at 473-75.
43 Id. at 475-80.
44 Id. at 481 (Owens, J., dissenting).
45 Id.
46 Markle Interests, LLC, 848 F.3d 635 (denying petition for rehearing en banc).
47 Id. at 636-37 (Jones, J., dissenting).
48 Id. at 637.
49 Weyerhaeuser, 138 S. Ct. 924.
should be given a wide berth in determining how to best protect endangered species.

a. The Designation Was Neither Arbitrary Nor Capricious

First, the Service argues that its designation of the Louisiana property as unoccupied critical habitat must be upheld unless it is found to be arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law, per the APA. The Service then argues that its designation was anything but. The Louisiana property was identified by the Service after peer reviewers criticized the initial proposed designation—which only included land in Mississippi—as inadequate. That led the Service to the Landowners’ property in Louisiana, which was said to be in the historical range of the frog. Although not perfect, the fact that the frog was reported to have been seen on the property many years ago convinced the Service that the property could be modified to conserve the frog and thus met the statutory requirements to serve as unoccupied critical habitat for the frog.

b. The Service’s Designation of the Louisiana Property Deserves Deference

That the Louisiana property is not a perfect habitat for the frog because it does not contain all the PCEs for the frog should not disqualify it from the designation; other courts have previously accepted this point in a variety of circumstances. The Service submits that to hold otherwise on these facts would be to reject the long-standing principle of deference to agency decision-making when it comes to areas within its expertise. And determining “habitat” for a species is a scientific question, not a legal one, as the Service sees it. To buttress that conclusion, the Service notes that its interpretation of “habitat” is consistent with the ESA’s purpose: to “provide a means whereby the ecosystems upon which endangered species and threatened species depend may be conserved.”

At bottom, Congress trusts the Service to protect endangered species, and it delegated power to the Service to carry out that important mission. The protection of endangered species “requires an expertise and attention to detail that exceeds the normal province of Congress.” Where even the Supreme Court has recognized that it is “beyond doubt that Congress intended

d. The Designation of Critical Habitat That Did Not Have All PCEs Markedly Differ from This Case

Second, the Landowners acknowledge that courts have approved the designation of critical habitat that did not include all PCEs for the endangered species, but those circumstances differed meaningfully from the instant case. In this case, the Louisiana designation is an unoccupied area unconnected from and unrelated to areas that provide the remaining essential features.
for the balance of a species’ life cycle. Cases that allowed for designation without all PCEs did not suffer from that deficiency. For example, in *Home Builders Association of Northern California v. U.S. Fish and Wildlife Service*, the Ninth Circuit addressed whether vernal pools and their immediate surrounding areas could be designated as occupied critical habitat for a species where the pools themselves contained *most* but not all of the primary constituent elements (PCEs) for the species. The Ninth Circuit held that since the two portions of the designation together provided all four PCEs necessary for the habitat, the ESA did not require that each portion of the designated area supply all of the PCEs independently of the other. In this case, on the other hand, the entirety of the Louisiana property, even when combined with immediately surrounding areas, does not include all three PCEs for the frog. The Service concedes this.

Simply put, an area cannot be “essential to a species conservation” if it is unlikely to contribute to that conservation at all. In *Home Builders*, it was likely that the habitat would contribute to the conservation of the species, especially in combination with an immediately adjacent area. That is not the case here. Notably, the Service recently recognized this logic in its proposal to amend its Regulations for Listing Species and Designating Critical Habitat, although it proposed that this change would only apply to future designations.65

B. Can Courts Review the Service’s Designation Decisions?

The second question presented by the case is whether the Service’s decision not to exclude the Louisiana property from the critical habitat designation is insulated from judicial review. The Service submits that Congress did not intend such decisions to be judicially reviewable. The Landowners argue that they should be able to show in court that the Service abused its discretion when it designated the Louisiana property. This question was not the primary focus of the parties’ briefing in the lower courts, so the Supreme Court’s decision to grant certiorari on it is especially interesting.

1. The Service’s Argument: The Text of the ESA and the Lack of Standards for Review Mean Designation Decisions Are Not Subject to Judicial Review

The ESA expressly authorizes judicial review of certain specified actions or failures to act by the Service and other federal agencies.66 And although the ESA does not explicitly provide for judicial review of other actions pursuant to the statute, the Supreme Court has held the Service’s application of the ESA’s substantive requirements is generally subject to judicial review under the APA.67 But the APA itself does not allow for judicial review “to the extent that . . . agency action is committed to agency discretion by law.”68

Such is the case here, according to the Service and the lower court. Section 4(b)(2) of the ESA provides no instruction concerning how the Service should exercise its discretion to either exclude or not exclude land from critical habitat designation. The Act simply provides that, when the Service thinks exclusion would be more beneficial than inclusion, the Service may exclude the area, assuming the exclusion would not lead to extinction. Without guidance beyond that minor caveat, the ESA does not identify how the Service should decide whether to exclude, and that lack of guidance makes the discretion exercised when choosing not to exclude unreviewable. Without a standard to review the decision, the decision is unreviewable. Ultimately, the economic impact the Landowners suffer because of the designation does not give rise to a requirement that the Service’s decision not to exclude the Louisiana property from designation be reviewable.

2. The Landowners’ Argument: ESA Amendments and Standards of Review To Be Found at Law Justify Reviewability of Designation Decisions

a. Congress Was Concerned About the Economic Impact of Designations Under the ESA, and Courts Should Be Able To Ensure Congress’s Concern Is Properly Addressed by the Service

The argument against judicial review of § 4(b)(2) decision-making under the ESA finds no support in the provision’s statutory or legislative history. The original ESA of 1973 lacked a definition of or process for designating critical habitat.69 To be sure, in 1978, the Court ruled in *Tennessee Valley Authority v. Hill*, that the ESA required the preservation of endangered species “whatever the cost.”70 But, in response, Congress amended the ESA to require the Service to consider economic and other non-biological impacts when designating critical habitats, and Congress authorized the Service to exclude property from designation on account of excessive costs.71 Thus, construing the APA’s “committed to agency discretion by law” bar to preclude review of decisions made under the Service’s § 4(b)(2) authority would thwart the ESA’s amended aim of “introducing some flexibility which will permit exemptions from the Act’s stringent requirements.”72 The courts should be able to review the Service’s decision not to exclude to see if it abused its discretion in failing to exempt the land from the ESA’s stringent requirements.

b. Meaningful Standards Exist for the Court To Apply When Reviewing the Service’s Designation Decisions

Moreover, there are standards that courts can apply in this and similar cases. As Justice Antonin Scalia put it when addressing a case involving § 701(a)(2) in his dissent in *Webster v. Doe*,

64 616 F.3d 983, 988 (9th Cir. 2010).
65 See 83 Fed. Reg. 35,193, 35,198 (July 25, 2018) (“In order for an unoccupied area to be considered essential, the Secretary must determine that there is a reasonable likelihood that the area will contribute to the conservation of the species.”). It is a mystery why the Service refuses to apply this new revision to past designations—a mystery the Justices of the Court will likely probe during the oral argument.
"[T]here is no governmental decision that is not subject to a fair number of legal constraints precise enough to be susceptible of judicial application—beginning with the fundamental constraint that the decision must be taken in order to further a public purpose rather than a purely private interest." Moreover, the Service itself identified a standard that courts could apply. In deciding not to exclude the Louisiana property, the Service explained it could not identify any “disproportionate costs” attendant to the designation. A court could review the facts of the case to determine whether the $34 million economic impact was a disproportionate cost where the critical habitat designation did not benefit the frog.

The Service may have wide discretion in assessing economic impact as compared to biological benefit, but there is scant evidence that Congress expected that discretion to be unfettered. Yet that is what the lower courts held, and it is what the Service seeks. The Supreme Court in recent years has repeatedly reversed lower court decisions that insulate agency decision-making from judicial review, and this case presents another opportunity for the Court to place limits on what agencies can do unchecked.

V. Conclusion

In the first case of its new term, the Supreme Court will consider the scope of the Service’s delegated powers under the ESA, and whether the Service’s exercise of those powers in the critical habitat designation for the dusky gopher frog is beyond judicial review. Given that there was no obvious circuit split supporting the grant of review, several of the Justices may think the Service went too far.


74 Compare Michigan v. EPA, 135 S. Ct. 2699, 2707 (2015) (“One would not say that it is . . . rational . . . to impose billions of dollars in economic costs in return for a few dollars in health or environmental benefits.”).

Will We Soon Have Clarity on Navigable Waters?:
How the Supreme Court’s October 2017 Term Set the Stage

By Tony Francois

Note from the Editor:
This article discusses the longstanding legal battle over the meaning of “navigable waters” in the Clean Water Act. It argues that several of the Supreme Court’s recent decisions have cleared a path for a final answer to this lingering question.

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.


The Clean Water Act (CWA) empowers the Environmental Protection Agency (EPA) to regulate “navigable waters.”1 Navigable waters are defined as “waters of the United States,”2 but that term is left undefined in the law. Prior to 2006, EPA defined waters of the United States to include all non-navigable tributaries to navigable waters, and all wetlands adjacent to (broadly defined as bordering, contiguous, or neighboring) either navigable waters or their non-navigable tributaries.3 These definitions were struck down by the Supreme Court in 2006 in Rapanos v. United States as exceeding the scope of the statutory term “navigable waters.” Following Rapanos, the EPA used informal guidance for several years to regulate tributaries and adjacent wetlands.4 But in 2015, the EPA promulgated a new rule defining navigable waters even more broadly than it had previously.5 This controversial rule was immediately challenged by landowners across the country who feared that streams and puddles on their land might soon invite federal government scrutiny and regulation, and consequently cause the value of their land to plummet. Some challenged the law after they were sentenced to fines and even jail time under criminal provisions of the CWA for polluting small bodies of water. Many of these challenges to the EPA’s 2015 definition have, until recently, been on hold in the lower courts awaiting jurisdictional decisions, possible changes to the regulation, and clarifications of law that could affect their outcome.

But after the Supreme Court’s October 2017 Term, the stage is set for a major decision on the geographic scope of the Clean Water Act. The Court decided three cases—one dealing with the CWA directly and two on related issues—that clear a path for such a decision by answering a threshold jurisdictional question, providing a useful framework for deciding vagueness cases, and shedding light on how lower courts should deal with fractured Supreme Court precedents. The Court held in National Association of Manufacturers v. Department of Defense (NAM v. DOD)6 that challenges to the EPA’s 2015 regulation should be brought in district courts rather than courts of appeal in the first instance, a necessary jurisdictional clarification.7 The Court’s immigration-related decision in Sessions v. Dimaya8 provided a development in void for vagueness law that may bear on how the Court decides the underlying substantive question under the CWA: What does navigable waters mean? The Court revisited a prior fractured

About the Author:
Tony Francois is a Senior Attorney with the Pacific Legal Foundation. He is counsel of record for Mr. Robertson in Robertson v. United States, and he represented several respondents supporting petitioners in NAM v. DOD; both cases are discussed in this article.

---

1 33 U.S.C. § 1311(a), § 1362(12).
3 33 C.F.R. §§ 328.3(a)(1), (a)(3), (a)(5), (a)(7), and 328.3(c) (2004).
decision in *Hughes v. United States* without ultimately clarifying the rule in *Marks v. United States*, leaving open the question of how lower courts should deal with fractured Supreme Court decisions like *Rapanos*.11

Meanwhile, the challenges to EPA’s 2015 navigable waters regulation are slowly working their way through the lower courts. A petition for certiorari will soon be presented to the Supreme Court in another CWA case turning on the meaning of navigable waters: *United States v. Robertson*.12

I. The Supreme Court Decided Which Federal Courts Should Hear the Navigable Waters Definition Cases

In January, the Court decided *NAM v. DOD*, which resolved a threshold procedural issue necessary for the ongoing litigation over whether EPA’s 2015 regulation defining “navigable waters” is legal.13 One of the CWA’s many technical provisions allocates alternative original jurisdiction over challenges to EPA actions in the federal district courts or federal circuit courts, depending on the type of EPA action being challenged.14 Over a hundred plaintiffs filed several lawsuits against EPA’s 2015 navigable waters definition in district courts around the country.15 Some of the cases were dismissed on the ground that jurisdiction lay in the circuit courts.16 In others, the district courts ruled that jurisdiction was proper.17 Most of the plaintiffs also filed protective petitions for review in the circuit courts, which were consolidated in the Sixth Circuit. But the National Association of Manufacturers did not file a protective petition. Instead, it intervened in the consolidated circuit court proceeding and moved to dismiss it for lack of jurisdiction, arguing that jurisdiction was proper in the district courts.18 The Sixth Circuit denied the motion and, in an unusual move, the Supreme Court granted certiorari to review that denial.19

*NAM v. DOD* holds that the district courts have original jurisdiction over the pending challenges to the 2015 navigable waters definition.20 In *NAM*, the Court took a textualist approach to determine whether the CWA vests original jurisdiction in the district or circuit courts.21 The CWA provides that suits should be filed originally in the federal circuit courts if they challenge EPA decisions that approve or promulgate an effluent limitation or other limitation under various provisions of the CWA, or that issue or deny any permit under 33 U.S.C. § 1342.22 The unanimous Court rejected the government’s atextual argument that the “practical effects” of the 2015 navigable waters definition effectively made it an “other limitation” by subjecting areas to permitting.23 The Court also refused to extend what the government called a “functional interpretive approach” found in *Crown Simpson Pulp Co. v. Castle*.24 *Crown Simpson* held that EPA vetoes of state-issued CWA permits were subject to immediate circuit court review because the veto is “functionally similar” to an EPA grant or denial of a permit, which is specifically subject to immediate circuit court review under the CWA.25 In *NAM*, the Court limited *Crown Simpson* to its facts and rejected the government’s call to extend a “functional interpretive approach” to other areas of the CWA.26 The Court also rejected the government’s appeals to judicial efficiency and national uniformity.27

Following *NAM v. DOD*, several of the cases challenging EPA’s 2015 navigable waters definition are now moving forward in district courts. Three of those courts have enjoined the regulation’s enforcement in 28 states.28 One of the injunctions is being reviewed in the Eleventh Circuit,29 but none has been resolved finally on the merits in district court. Given how broadly EPA defined “navigable waters” and the Supreme Court’s ongoing interest in the issue, it seems certain that the Court will review these cases or otherwise address the question in similar litigation.30

---

21 See id. at 628-30 (interpreting “effluent limitation or other limitation”), id. at 631 (interpreting “issuing or denying any permit under section 1342”).


23 Id. at 630-31.


25 *Crown Simpson*, 445 U.S. at 196 (referring to the CWA provision found at 33 U.S.C. § 1369(b)(1)(F)).

26 *NAM*, 138 S. Ct. at 632.

27 Id. at 633-34.


II. LAWSUITS CHALLENGING THE 2015 NAVIGABLE WATERS REGULATION ARE UNLIKELY TO BECOME MOOT

The fact that EPA is rewriting its 2015 navigable waters definition probably does not lessen the likelihood of eventual Supreme Court review of that regulation.31 The rewrite is not complete, and while EPA has predicted that it will be complete by the end of 2018, it remains uncertain whether that prediction will turn out to be accurate. As of this writing, a new proposed definition has not been published.32

The complexity of EPA’s ongoing regulatory work further decreases the likelihood that the suits challenging the 2015 regulation will be mooted in the immediate future. In February 2018, EPA adopted what it called the Applicability Date Rule.33 This rule purports to advance the date on which the 2015 Water Definition “is applicable” to February 2020, but without changing the effective date of the regulation.34 Environmentalists and states have sued the EPA arguing that the Applicability Date Rule is invalid.35 On August 16, 2018, the U.S. District Court for the District of South Carolina enjoined the Applicability Date Rule nationwide.36 Thus far, EPA’s gambit has done nothing to moot the pending lawsuits against the 2015 regulation.37 Two of the three injunctions against the 2015 definition were entered months after the adoption of the Applicability Date Rule.38

Meanwhile, EPA is preparing two separate rulemakings, one to repeal the 2015 definition (the Repeal Rule) and another to adopt a new definition (the Replacement Rule).39 The Repeal Rule is expected to be issued prior to the Replacement Rule,40 and it is almost certain that environmental activists and some states will sue to invalidate the Repeal Rule.41 The Repeal Rule proposes to rescind the 2015 regulation, which was in effect throughout most of the nation from August 28 to October 5, 2015, and adopt the previous regulations without substantively analyzing them. Since this re-adoption of the pre-2015 regulations without substantive comment was a legal flaw in the Applicability Date Rule it seems likely that the Repeal Rule will be at least temporarily enjoined, leaving the 2015 navigable waters definition and the suits against it in effect. The same environmental activists and states have also promised to sue over the Replacement Rule when it is adopted, with similar prospects for an injunction. This would leave the lawsuits against the 2015 regulation unmooted, despite the Trump Administration’s best efforts to repeal and replace the Obama Administration’s rule.

If a new Trump Administration regulation defining navigable waters goes into effect and survives legal challenge, the Supreme Court would likely address the definition of navigable waters in environmental plaintiff challenges to EPA’s Replacement Rule.42

III. THE SUPREME COURT LEFT FOR ANOTHER DAY A NEEDED CLARIFICATION OF HOW TO INTERPRET ITS FRACTURED OPINIONS

The Court’s resolution of another case sets the stage for the Court to revisit its 2006 fractured decision in Rapanos on the definition of navigable waters. In June, the Court decided Hughes v. United States, holding that criminal defendants who are sentenced under certain types of plea agreements are eligible for resentencing if the Sentencing Guidelines were revised and their sentences were based on the revised Guidelines.43

What does that have to do with the Clean Water Act?44 Hughes was granted to resolve a circuit split over how to apply the Supreme Court’s 2010 fractured decision in Freeman v. United States.45 Justice Kennedy, writing for a four-Justice plurality in Freeman, took the view that defendants who had entered plea agreements were eligible for resentencing if the judge had relied on the subsequently revised Guidelines in adopting the

1812 n.1 (2016) (noting adoption of 2015 navigable waters rule and its nationwide stay by the Sixth Circuit); NAM, 138 S. Ct. at 625 (“In 2015, responding to repeated calls for a more precise definition of “waters of the United States,” the agencies jointly promulgated the navigable waters regulation).


34 Id. at 5201.


37 EPA argued against an injunction in one of the pending lawsuits that the Applicability Date Rule weighed against enjoining the 2015 navigable waters regulation. See Texas v. EPA, S.D. Tex. No. 3:15-cv-00162, Dkt # 101 at 2.


42 See id.

43 138 S. Ct. at 1774-77.

44 Two amicus briefs filed in Hughes argued that the case was of critical importance to the Clean Water Act. See Brief Amici Curiae of Chantell and Michael Sackett and Duarte Nursery, Inc., in Support of Petitioner, Hughes v. United States, No. 17-155, 2018 WL 620239 (U.S. Jan. 25, 2018) (arguing that the plurality is the holding of Rapanos under Marks), and Brief Amici Curiae for Agricultural, Building, Forestry, Livestock, Manufacturing, Mining, and Petroleum Business Interests in Support of Petitioner, Hughes v. United States, No. 17-155, 2018 WL 620238 (U.S. Jan. 29, 2018) (arguing that neither the plurality nor the concurrence is the holding of Rapanos under Marks).

45 564 U.S. 522.
Kennedy’s view to form a majority on the legal issue. As a result, abandoned her concurrence and joined in Justice Freeman. Under States determine the case’s holding, if any, under Marks v. United States. Under Marks, the holding of a fractured decision is the opinion of those Justices who concurred in the judgment on the narrowest grounds. Despite the apparent simplicity of this test, circuit courts have been bedeviled in their efforts to apply Marks consistently.

The Supreme Court granted certiorari in Hughes on two questions involving how the lower courts should apply Marks, and this drew the case within the ambit of CWA jurisprudence: any clarification of Marks in Hughes could have been applicable also to a case applying the Court’s fractured Rapanos precedent. But the Court decided Hughes without addressing Marks by eliminating the original split in Freeman. Justice Sotomayor abandoned her Freeman concurrence and joined in Justice Kennedy’s view to form a majority on the legal issue. As a result, the Court had no need to say anything substantive about Marks.

The Court’s failure to resolve the questions related to Marks in Hughes leaves those questions open for another day. The amicus brief in Hughes on Marks and Rapanos highlighted the CWA as an important area in which the Court’s clarification, either of how to apply Marks or of the underlying substantive question, is badly needed.

IV. The Court Decided an Important Void for Vagueness Case That May Bear On Its Ultimate View of Navigable Waters

Now that they know which courts have jurisdiction over their lawsuits, the NAM litigants (and those similarly situated) can get to the merits: What are “navigable waters” under the Clean Water Act, and does EPA’s 2015 navigable waters definition fit within or exceed that meaning? The Court’s March 2018 decision in Sessions v. Dimaya held a provision of the Immigration and Nationality Act (INA) unconstitutionally vague, and it provides possible insight into the answer to the first part of that question. One issue that has dogged the effort to determine what counts as navigable waters (given that some areas so designated are neither “navigable” nor even “water” for much of each year) is whether the statutory term is unconstitutionally vague. Since the Court’s 2006 fractured decision in Rapanos v. United States, the lower courts have largely adopted Justice Kennedy’s lone concurrence, which holds that “navigable waters” are determined through a case-by-case inquiry for a “significant nexus” between the wetlands or tributaries at issue and downstream traditionally navigable waters. “Significant nexus” is determined across three separate criteria—physical, chemical, and biological—using highly subjective factors. In practice, this interpretation of navigable waters frequently boils down to “I know it when I see it” subjective determinations by EPA or Army Corps field staff.

Since Rapanos, members of the Supreme Court have observed that the “significant nexus” interpretation of “navigable waters” leaves regulated citizens with little or nothing to go on in figuring out if their property or activities are subject to the Act.

---

54 530 U.S. 188.
55 Id. at 551 (Roberts, C.J., dissenting).
56 138 S. Ct. at 1771.
57 138 S. Ct. at 1204.
58 See United States v. Johnson, 467 F.3d 56 (1st Cir. 2006); United States v. Gundluff, 555 F.3d 200 (6th Cir. 2009); Gibson v. American Cyanamid Co., 760 F.3d 600 (7th Cir. 2014); United States v. Bailey, 571 F.3d 791 (8th Cir. 2009); Northern California River Watch v. City of Healdsburg, 496 F.3d 993 (9th Cir. 2007).
59 Rapanos, 547 U.S. at 758 (Kennedy, J., concurring).
60 The practice is so subjective and staff-dependent that the government takes the position that a formally adopted Jurisdictional Determination, which is the Army Corps’ final word on whether a given feature is a “navigable water” under the Act, is nonetheless nonbinding on the EPA in its exercise of its parallel enforcement authority. See Hawkes, 136 S. Ct. at 1817 (Kennedy, J., concurring).
61 See, e.g., Rapanos, 547 U.S. at 758 (Roberts, C.J., concurring) (“Lower courts and regulated entities will now have to feel their way on a case-by-case basis.”); Sackett, 556 U.S. at 124 (“The Sacketts are interested parties feeling their way.”); id. at 132 (Alito, J., concurring) (“The reach of the Clean Water Act is notoriously unclear.”); id. at 133 (the phrase “waters of the United States” is “not a term of art with a known meaning” and is “hopelessly indeterminate”); Hawkes, 136 S. Ct. at 1812 (“It is often difficult to determine whether a particular piece of property contains waters of the United States.”); id. at 1816-17 (Kennedy, Alito, Thomas, JJ., concurring) (“[T]he reach and systematic consequences of the Clean Water Act remain a cause for concern.”) (quoting Sackett, 556 U.S. at 132 (Alito, J., concurring)); NAM v. DOD, 138 S. Ct. at, 625 (“In decades past, the EPA and Corps . . . have struggled to define and apply that statutory term.”). See also Hawkes v. Army Corps, 782 F.3d 994, 1003 (8th Cir. 2015) (Kelly, J., concurring) (“[T]he Court in Sackett was concerned with just how difficult and confusing it can be for a landowner to predict whether or not his or her land falls within CWA jurisdiction . . . . This is a unique aspect of the CWA; most laws do not require the hiring of expert consultants to determine if they even apply to your property.”); Orchard Hill Building Company v. Army Corps, 893 F.3d 1017, 1025 (7th Cir. 2018) (“Justice Kennedy did not define ‘similarly situated’—a broad and ambiguous term . . . .”)
During oral argument in 2016’s U.S. Army Corps of Engineers v. Hawkes, Justice Kennedy posed the following question:

Well, I think—I think underlying Justice Kagan’s question is that the Clean Water Act is unique in both being quite vague in its reach, arguably unconstitutionally vague, and certainly harsh in the civil and criminal sanctions it puts into practice. What’s the closest analogous statute that gives the affected party so little guidance at the front end?62

Dimaya held that 18 U.S.C. § 16(b), defining “crime of violence” for purposes of the INA, is void for vagueness.63 The Court’s analysis rested on the statute’s use of two terms: “by its nature” (as applied to the noun “felony”) and “substantial risk” (that physical force would be used in committing the crime). Both terms require an interpreting court to decide, without any standards, what crimes fall within the definition.64 Relying heavily on its prior decision in Johnson v. United States, the Court noted that applying the “by its nature” provision requires a court to determine the “idealized ordinary case” of a given offense.65 And that exercise yields no clear answer; it depends entirely on a given judge’s opinion of what the essential nature or “platonic form” of a given crime involves.66 Secondly, this indeterminacy is compounded by the requirement that the judge then determine whether the platonic form of a crime poses some threshold level of risk—a “substantial risk”—of violence.67 The Court grants the constitutionality of applying a “substantial risk” standard, standing alone, to a defendant’s conduct. It is the combination of the need to posit an idealized version of a crime with the question whether the idealized form poses a threshold risk level which crosses the line into vagueness.68

The same analytical approach is applicable to Justice Kennedy’s interpretation in Rapanos of navigable waters under the CWA. As with the statute struck down in Dimaya, Justice Kennedy’s Rapanos concurrence interprets the CWA term “navigable waters” to require two interacting determinations, one involving an idealized or otherwise undefinable condition (“wetlands . . . in combination with similarly situated lands in the region”), and the second overlaying a threshold relationship (“significantly affect” traditionally navigable waters).69

The “similarly situated within the region” provision requires a judge to make two idealized determinations: what two or more wetlands are “similarly situated” to each other, and what is “the region” within which those wetlands’ situation must be similar. As interpreted by Justice Kennedy, “navigable waters” offers no guidance to answer either of these questions.70 Wetlands can be similar in any number of ways: location, size, plant communities, length of inundation, type of connection to other features, animal communities that use or rely on them, soil types, etc.71 They may, at the same time, be similar in some of these aspects and dissimilar in others. How is a judge (or regulated party, agency staff, administrative law judge, or citizen suit plaintiff or defendant) to determine whether any two or more wetlands are similarly situated to a degree that satisfies Justice Kennedy’s interpretation of the CWA?72 Nor is the platonic form of “the region” any more determinate. How large is a region? And how are its borders defined? If by watershed, how large a part of the watershed? The portion in which the similarly situated wetlands appear, or the entire watershed of the applicable traditionally navigable water? The larger the region (whether defined by a watershed or some other geographic concept), the more indeterminate “similarly situated” becomes.

In this respect, Justice Kennedy’s “similarly situated within the region” interpretation of “navigable waters” is even less knowable for the regulated citizen or enforcement personnel than the “ordinary case” of any given crime under 18 U.S.C. § 16(b). Justice Kennedy’s “significant nexus” requires two abstract determinations—“similarly situated” and “the region”—that entirely depend upon the subjective judgment of the reviewing court or enforcing agency staff, whereas the “ordinary case” of a crime only requires one such imaginative abstraction.

And exactly as in Dimaya and Johnson, this abstracted concept of similarly situated wetlands in a region is overlaid by an equally problematic significance threshold: a significant nexus with downstream traditionally navigable waters.73 The combination of the idealized “similarly situated within the region” wetland combination that also “significantly affects” downstream, boat-floating, commerce-supporting rivers and lakes renders Justice Kennedy’s reading of “navigable waters” hopelessly vague and far short of constitutional muster, as he indeed intimated during the Hawkes oral argument.

V. The Supreme Court Could Clear Up Navigable Waters in Robertson v. United States

In sum, Hughes leaves unresolved questions as to the application of Marks v. United States, which could be resolved in the context of a case addressing Rapanos, either through clarification of the Marks framework or by replacing the underlying fractured decision with a new majority opinion. And Dimaya offers a robust analytical framework demonstrating that

---

63 138 S. Ct. at 1210.
64 Id. at 1213-14.
65 138 S. Ct. at 1214 (quoting Johnson, 135 S. Ct. 2551, 2557 (2015)).
66 138 S. Ct. at 1214; id. at 1231-32 (Gorsuch, J., concurring).
67 Id. at 1214.
68 Id. (quoting Johnson, 135 S. Ct. at 2561).
69 547 U.S. at 780 (Kennedy, J., concurring).
70 Id. (“wetlands, either alone or in combination with similarly situated lands in the region”).
72 The Supreme Court has interpreted or applied the term “similarly situated” in a variety of contexts, suggesting that while its meaning varies based on context, it suggests similarity in aspects or function rather than merely being nearby each other. See, e.g., Hoffmann-La Roche Inc. v. Sperling, 493 U.S. 165 (1989) (applying notice provision under Age Discrimination in Employment Act for “similarly situated” employees).
73 See Dimaya, 138 S. Ct. at 1215-16.
EPA’s 2015 navigable waters definition, to the extent it is based on Justice Kennedy’s *Rapanos* concurrence, is invalid because that reading of the CWA is unconstitutionally vague.74

In addition to the litigation over the 2015 navigable waters definition, a petition has been filed in the Supreme Court in another CWA case. On July 10, 2018, the Ninth Circuit denied rehearing in a criminal appeal that clearly frames both the *Marks* and the void for vagueness issues: *United States v. Robertson*.75

The federal government prosecuted Mr. Robertson under the CWA76 for his impacts to a 12-inch-wide, 18-inch-deep channel77 carrying 2–3 garden hoses worth of flow,78 several miles from the nearest actually navigable river in rural Montana. He was ultimately imprisoned for 18 months. One of Mr. Robertson’s defenses is that the CWA’s phrase “navigable waters” is void for vagueness.79 The Ninth Circuit rejected that defense on the ground that Justice Kennedy’s concurring opinion in *Rapanos* has been held by the Ninth Circuit to be the controlling definition of “navigable waters,” which alone provides adequate notice of the law’s requirements.80 But the Ninth Circuit said nothing about whether “navigable waters” itself, as interpreted by Justice Kennedy, is void for vagueness.

In applying the *Marks* framework to the *Rapanos* decision to decide *Robertson*, the Ninth Circuit expressly held that circuit courts may use dissenting opinions to fashion a holding for fractured Supreme Court decisions.81 The Ninth Circuit thus created a circuit split with the Seventh and DC Circuits on that precise question,82 raising yet another question about *Marks* that warrants Supreme Court clarification.

Mr. Robertson’s cert petition offers the Court a vehicle to apply the *Dimaya* framework to “navigable waters,” as interpreted by both by the plurality and by Justice Kennedy, and to address whether Justice Kennedy’s reading of navigable waters is even the controlling rule of law from *Rapanos* under *Marks*. Given the complexity of both the administrative rulemaking at EPA to adopt regulations defining navigable waters and the ongoing litigation over that process, it could be more efficient for the Supreme Court to resolve these questions in the context of Mr. Robertson’s appeal from his criminal conviction. Such a decision could provide much needed guidance, for example, to EPA in its ongoing efforts to write the new definition of navigable waters.

A clear majority decision on the meaning of navigable waters could also end the interminable political battle over the scope of the CWA, in which the prevailing political faction uses its control of EPA and the Army to revise guidance and regulations in order to expand or contract the meaning of navigable waters to suit its constituents. This process has replaced the rule of law with naked partisanship. The Supreme Court should end the scrum by clearly and definitively ruling on the meaning of the term Congress actually enacted, restoring the rule of law to this important area.

74 Where there are multiple reasonable interpretations of a statute, the rule of lenity interacts with the void for vagueness doctrine to limit criminal statutes to activity clearly covered. See *United States v. Lanier*, 520 U.S. 259, 266 (1997) (discussing doctrine, reciting elements, and citing sources) (citing *Liparota v. United States*, 471 U.S. 419, 427 (1985), and others). This rule would require a preference for Justice Scalia’s plurality opinion in *Rapanos* to the extent it interprets “navigable waters” without violating the Due Process fair notice requirement.

75 875 F.3d 1281, *rehearing and rehearing en banc den.* July 10, 2018. The author is counsel of record for Mr. Robertson.

76 *Id.* at 1286.

77 *See United States v. Robertson*, 9th Cir. Docket No. 16-30178, Excerpts of Record, Vol. 4 Dkt # 16-4 at 227:10-11.

78 *Id.* Excerpts of Record, Vol. 11, Dkt # 16-11 at 42:5-7.

79 875 F.3d at 1292.

80 *Id.* at 1293.

81 *Id.* at 1291.

82 Compare *Robertson*, 875 F.3d at 1292 (forming *Marks* holding by combining concurrence with dissent), with *Gibson v. American Cyanamid Co.*, 760 F.3d 600, 621 (7th Cir. 2014) (dissects may not be used to form a holding under *Marks*); see also *King v. Palmer*, 950 F.2d 771, 783 (D.C. Cir. 1991) (en banc) (“[W]e do not think we are free to combine a dissent with a concurrence to form a *Marks* majority.”).
COUNTING TO TWO THIRDS: HOW CLOSE ARE WE TO A CONVENTION FOR PROPOSING AMENDMENTS TO THE CONSTITUTION?

By Robert G. Natelson

Note from the Editor:

This article argues that, in aggregating applications from states to call a convention for proposing amendments under Article V of the U.S. Constitution, Congress should count plenary (unlimited) applications toward a limited-subject convention.

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.


About the Author:

Rob Natelson is Professor of Law at The University of Montana (ret.), Senior Fellow in Constitutional Jurisprudence at the Independence Institute in Denver, Colorado, and Director of the Independence Institute’s Article V Information Center. His work has been relied on repeatedly by U.S. Supreme Court justices and parties and has been cited by the highest courts of 15 states.

Article V of the United States Constitution provides that when two thirds (currently 34) of the state legislatures apply, “Congress . . . shall call a Convention for proposing Amendments.” To determine whether its duty to call a convention has been triggered, Congress must count applications from states; this practice sometimes is referred to as “aggregating” applications. This paper addresses the almost unexamined question of whether applications for a convention unlimited as to topic (“plenary applications”) should be aggregated with those for a convention limited to one or more subjects.

Congress may face this issue very soon. At least 27 state legislatures have valid applications outstanding for a convention to propose a balanced budget amendment (BBA). At least six states without BBA applications have outstanding applications calling for a plenary convention. Thus, if aggregation is called for, 33 of the 34 applications needed for Congress to call a convention likely exist.

After consideration of the language of Article V, case law, historical practice, and other factors, this paper concludes that Congress should add existing plenary applications to the BBA
total, and that it should call a BBA convention if and when the aggregated total reaches 34.

I. Basic Principles

Article V provides that, to become part of the Constitution, an amendment must be ratified either by (1) three fourths of the state legislatures or (2) conventions in three fourths of the states. Congress chooses between the legislative and convention ratification methods. However, before an amendment may be ratified, it first must be duly proposed. Article V itemizes two permissible methods of proposal: (1) by a two thirds vote of both houses of Congress or (2) by “a Convention for proposing Amendments.” This paper focuses on the latter method, which the framers designed as a way of proposing amendments without the consent of Congress.

Article V does not delinate expressly the composition and nature of a convention for proposing amendments, and such a convention has never been held. For this reason, commentators, particularly those who oppose a convention, have long complained that Article V provides insufficient guidance on the subject. But the brevity of Article V is consistent with the drafting of the Constitution generally. The Framers sought to keep the document short by outlining the basics and leaving to readers the task of supplementing the text from contemporaneous law and circumstances. For example, Article I, Section 9, Clause 2 states that “The privilege of the writ of habeas corpus shall not be suspended . . . .” It does not explain what a writ of habeas corpus is, what it contains, how it is issued, or the traditional rules regarding suspension. Readers are expected to identify those facts for themselves. In this respect, Article V is no different.

Recent scholarly investigations into Article V have placed in the public domain the information necessary for understanding the Article V convention process. For example, both Founding-Era evidence and the Supreme Court inform us that a convention for proposing amendments is a kind of “convention of the states”—also called a “convention of the states.” This characterization has the effect of clarifying basic convention protocols, because the protocols of such conventions were standardized long before the Constitution was drafted: The Constitutional Convention of 1787 was a convention of the states, and it had over thirty predecessors. In fact, many of the delegates at the Constitutional Convention were veterans of one or more previous interstate gatherings.

Moreover, the protocols have not changed significantly since the Founding. Conventions of states met in Hartford, Connecticut (1814); Nashville, Tennessee (1850); Washington, D.C. (1861), Montgomery, Alabama (1861) St. Louis, Missouri (1889); Santa Fe, New Mexico and three other cities (1922); in various locations from 1946 to 1949; and in Phoenix, Arizona (2017). Although the specific rules for each meeting differed somewhat, the basic protocols remained roughly similar. Most interstate conventions, both before and after the ratification of the U.S. Constitution, have been regional or “partial” conventions to which colonies or states from only a single region of the country were invited. At least eight have been general conventions—that is, gatherings to which colonies or states from all regions were invited. An Article V convention for proposing amendments would be general, but there are no significant protocol differences between partial and general conventions. Those protocols determine such matters as the scope of a convention call, how commissioners are instructed, and how rules are adopted.

Article V does not outline these details because they were so well known to the founding generation that there was no need to repeat them. Article V is more specific only in a few instances where clarification was necessary. In view of the wealth of history surrounding Article V, the courts appropriately defer to that history. The Supreme Court and other judicial tribunals have decided nearly fifty reported Article V cases, and they

---

3 U.S. Const. art. V.

4 E.g., Laurence H. Tribe, Issues Raised by Requesting Congress to Call a Constitutional Convention to Propose a Balanced Budget Amendment, 10 Paq. L.J. 627, 632 (1979) (calling the Constitution’s convention wording “strikingly vague”).


7 Convention of the States, supra note 2.

8 Smith v. Union Bank, 30 U.S. 518, 528 (1831) (referring to a convention for proposing amendments as a “convention of the states”).

9 The constitutional term “convention” is probably the most common designation, but at various times, they also have been known as interstate congresses, committees, and commissions. See generally Founding-Era Conventions, supra note 2; Robert G. Natelson, List of Conventions of States in American History, http://articlevinfocenter.com/list-conventions-states-colonies-american-history/.

10 Founding-Era Conventions, supra note 2, at 691-710 (identifying attendees at the Constitutional Convention and prior Founding-Era conventions, initially listed by alphabetical order for each attendee, and then grouped by state).


12 For example, at all of these conclaves states enjoyed equal voting power. Specifically, at every convention except St. Louis (1889), each state had one vote. At St. Louis, each state had eight votes. Robert G. Natelson, Newly Rediscovered: The 1889 St. Louis Convention of States, http://articlevinfocenter.com/newly-rediscovered-1889-st-louis-convention-states/.

13 Id. The general conventions were Albany (1754), New York City (1765 and 1774), Annapolis (1786), Philadelphia (1780 and 1787), Washington, D.C. (1861), and Phoenix (2017). Id.

14 The standard protocols originally were based on international practice. Caplan, supra note 2, at 95-96.

15 Founding-Era Conventions, supra note 2, at 686-90.

16 Id. at 689-90.

17 See GUIDE, supra note 2, at 12-13 for a table of cases.
have repeatedly consulted history to clarify the article’s words and procedures.  

II. Definitions of Terms

When the Constitution was adopted, an application was an address from one person or entity to another.  

It was thus a very broad term, and it could include communications among equals or between superiors and inferiors. An application could be an invitation, a request, a delegation, or an order.  

or between superiors and inferiors. An application could be an broad term, and it could include communications among equals  

or between superiors and inferiors. An application could be an broad term, and it could include communications among equals  

One kind of application was a convention call. This was an official invitation, often called a “circular letter,” sent to all or some states to meet at a particular time and initial place to discuss topics itemized in the call. Most calls were issued by individual states; others came from Congress or prior conventions. Calls were limited to time, initial place, and topic. Additional material, on the rare occasions when it was included, was prec atory.

Another kind of application, which might also be communicated by circular letter, encouraged the recipient to call or support a convention. Thus, a 1783 request from the Massachusetts legislature to the Confederation Congress asking it to call a convention was styled an “application.” To similar effect was the report of the 1786 Annapolis convention suggesting to the states that they meet in Philadelphia the following year, and the circular letter of July 26, 1788 issued by the New York ratifying convention urging another convention to consider amendments to the 1787 Constitution.

Calls and other convention applications almost invariably informed the recipients of the subjects for which the convention was sought. They almost never said merely, “let’s meet.” Rather, they said, “let’s meet to discuss trade issues”—or defense issues, or financial issues, or some specified combination. Calls and applications specifying different topics were understood to require different conventions. In 1786, one convention call invited all states to discuss trade issues while another invited some states to discuss navigation issues. There was no move to aggregate the two into a single meeting to discuss both.

Another class of applications not mentioned in Article V but inherent in any convention of states are those directed by principals to their agents—that is, from state legislatures to their representatives. In this class are commissions (also called credentials) whereby legislatures designate their commissioners. A commission is much like a power of attorney in that it names and empowers one or more agents and defines the scope of their authority. Each commissioner presents his or her commission to the convention before he or she may be seated. Closely related are instructions. As their name indicates, they contain more detailed directions from the appointing authority. Historically, commissions usually have been public documents while separate instructions often have been secret.

Article V refines to a certain extent how calls and other initial applications operate in the amendment context: Article V provides that state legislatures may apply to Congress, and when two thirds of them have done so, Congress must call an amendments convention. This enables state legislatures to promote amendments in a way that forestalls congressional veto. The congressional role in the convention process is mandatory and limited—ministerial rather than discretionary. Congress acts as a convenient common agent for the state legislatures. It follows necessarily that Congress’s function as the calling agent does not entitle it to alter traditional rules. Nothing in the Constitution supports the notion that Congress can expand its role to include, for example, dictating how commissioners are selected or what convention rules must be.

One last point pertains to terminology: Some commentators have referred to an unlimited convention as a “general convention.” This usage is incorrect. A general convention is a conclave to which states from all regions of the country are invited—as

18 Id. at 26, n.54 (collecting cases relying on history).
20 Id. Thus, a call sometimes was labeled an application. E.g., 1 Public Records of the State of Connecticut 589 (Charley Hoadley ed., 1894).
21 Founding-Era Conventions, supra note 2 (identifying the calling entities for major conventions held before 1788).
22 See generally id.
23 Id. at 667.
25 2 Jonathan Elliot, The Debates in the Several State Conventions on the Adoption of the Federal Constitution 413-44 (1836) (communicating with the governors of other states and urging them to support another convention).
26 See generally Founding-Era Conventions, supra note 2.
27 Id. at 668-72 (discussing the Annapolis Convention of 1786 and a proposed “Navigation Convention”).
28 See, e.g., The Federalist No. 40 (James Madison) (“The powers of the convention ought, in strictness, to be determined by an inspection of the commissions given to the members by their respective constituents.”); see also Caplan, supra note 2, at 97.
31 Caplan, supra note 2, at 94.
32 Professor Charles Black of Yale Law School may have originated the notion that Congress can control convention protocols. Charles L. Black, Jr., The Proposed Amendment of Article V: A Threatened Disaster, 72 Yale L.J. 958, 964-65 (1963). To support this view, he relied on the Necessary and Proper Clause. However, that Clause does not apply to the amendment process. See Guide, supra note 2, at 48-52. As the title suggests, Black’s article was polemical rather than scholarly in nature.
33 Professor Black seems responsible for this error as well, Charles L. Black, Jr., Amending the Constitution: A Letter to a Congressman, 82 Yale L.J. 189, 198 (1972), although others have repeated it.
opposed to a partial or regional gathering. A convention for proposing amendments is necessarily general, but may be limited or unlimited as to topic. If unlimited as to topic, it should be referred to as unlimited, open, or plenary.  

III. Article V Applications Must be Aggregated by Subject Matter

Only about twenty state legislative applications under Article V have been plenary—that is, seeking an unlimited or plenary convention. The other applications have sought conventions to consider amendments on one or more designated subjects. Article V does not provide expressly that the required two thirds of applications must address the same or overlapping subjects. This has led some to argue that because there have been far more than 34 applications, a call for a plenary convention is already mandatory. In other words, all valid applications must be aggregated with all other valid applications to yield a plenary result.

Three aspects of this argument render it unlikely of congressional or judicial acceptance. Most fundamentally, perhaps, it conflicts with the dictates of common sense: If 12 legislatures seek a convention to consider term limits, 12 seek a convention to consider a BBA, and 12 apply for a convention to consider campaign finance reform, it does not follow that 36 legislatures want a convention to consider everything, or all three topics, or any one of them. Further, this argument conflicts with Article V’s background history. In the Founders’ experience, convention calls and pre-call requests almost invariably designated one or more subjects and promoted a convention to address those subjects. Without prior agreement, states did not combine unrelated applications in a single convention.

Third, the argument conflicts with post-constitutional understanding. Consider by way of illustration the situation in the year 1911. At that time, there were 46 states, so 31 were needed to call a convention. Twenty-nine states had issued applications for a convention to propose direct election of U.S. Senators. Thirteen states had outstanding applications for a convention to propose a ban on polygamy. Subtracting states with applications on both subjects leaves 32—one state more than the required two thirds. Yet there is no evidence of widespread (or, indeed, any) contentions that direct election applications should be aggregated with anti-polygamy applications to force a convention. Not surprisingly, therefore, most commentators have concluded, or at least assumed, that for applications to aggregate they should overlap to some extent. This certainly has been the tacit assumption of Congress.

But to what extent must they overlap? Surely they need not be exact copies of each other. Founding-Era conventions met even though applications and instructions differed. In my 2016 treatise on the convention process, I addressed the question of how much coincidence is required. I listed four aggregation scenarios, as follows:

1. All applications seem to address the same subject, but restrictive wording in some renders them inherently inconsistent with others.
2. Some applications prescribe a convention addressing Subject A (e.g., a balanced budget amendment) while others prescribe a convention addressing both Subject A and unrelated Subject B (e.g., term limits).
3. Some applications prescribe a convention addressing Subject A (e.g., a balanced budget amendment) while others demand one addressing Subject X, where Subject X encompasses Subject A (e.g., fiscal restraints on the federal government).
4. Some applications prescribe a convention addressing Subject A and others call for a convention unlimited as to topic.

The treatise examined the first three scenarios in light of history, including the Founders’ own interpretive methods, and concluded that applications in the first two situations did not aggregate, but those in the third situation did. Because a full analysis of #4 would have consumed a disproportionate share of the treatise, I merely listed some arguments for both conclusions and suggested

---

34 Another possible kind of convention is “plenipotentiary.” This term is best reserved for conclaves meeting outside constitutional restraints—i.e., those that James Madison described as reverting to “first principles.” James Madison to G.L. Turberville, Nov. 2, 1788, 5 The Writings of James Madison 298-300 (Gaillard Hunt ed., 1904). By contrast, a convention for proposing amendments, even a plenary one, is limited to proposing amendments to the existing Constitution, and is subject to “the forms of the Constitution.” As explained below, states sometimes have sent commissioners with plenipotentiary powers to more limited conventions.

35 See The Article V Library, article5library.org. As of this writing, the Article V Library is the best and most reliable source for applications. There is at least one other website devoted to applications (http://foavc.org/), but it contains notable errors, including aggregating applications that do not overlap as to topic. A list of applications and rescissions kept by the Clerk of the U.S. House of Representatives at http://clerk.house.gov/legislative/memorials.aspx is incomplete and dates back only to 1960.

36 The most distinguished writer to urge this position is Michael Stokes Paulsen. See Paulsen, supra note 2, at 746-47. Professor Paulsen argued that an application conditioned on set topics was void, but that listing a particular change as its purpose should count toward a plenary convention. Professor Paulsen wrote in 1993, well before most of Article V’s defining history was recovered, although five years earlier Russell Caplan had documented the Founding-Era expectation that most applications would be limited. Caplan, supra note 2, at 95-99.

37 Founding-Era Conventions, supra note 2, at 668-72 (discussing the Annapolis Convention of 1786 and a proposed “Navigation Convention,” with no suggestion that the two be aggregated).

38 For lists of applications by date and subject matter, see the Article V Library, article5library.org.

39 E.g., Caplan, supra note 2, at 105 (“Twenty-four applications for a balanced-budget convention, and ten for a convention to consider school busing, will impose no duty on Congress”); See also Res. supra note 2, at 89 (“It seems obvious that if seventeen States apply for a convention to consider anti-abortion amendments, for instance, and seventeen others apply for a convention on a balanced budget amendment, the requisite consensus does not exist.”).

40 Cf. id. at 107 & 108.

41 Guide, supra note 2, at 55.

42 Id. at 56-58.
that an application’s specific wording might be helpful in weighing whether the application should be aggregated.\(^4\) The present paper examines the question more thoroughly. In doing so, we need not refer to hypothetical Subjects A, B, and X, because current events provide us with a real-life situation. Should BBA and plenary applications be aggregated together?

IV. Why Older Unrescinded Applications are Still Valid

Before proceeding further, I should explain why the extant (unrescinded) BBA and plenary applications remain valid even though several BBA applications are over 40 years old and the plenary applications are even older. Why have they not lapsed with passage of time?

During the 20th century, there was considerable discussion of this “staleness” question.\(^4\) Even the Supreme Court speculated on the staleness question as it pertains to ratifications of amendments,\(^4\) although no court has ever ruled on it. The intervening years have fairly well resolved the question for us: Unless expressly time-limited, applications remain in effect until formally rescinded. There are at least five reasons for so concluding.

First: Legislative actions normally do not lapse due to the mere passage of time. If their text does not limit their duration, they remain in effect until repealed, even if they become outdated. Nothing in constitutional history or usage suggests that Article V legislative resolutions comprise an idiosyncratic exception.

Second: The Twenty-Seventh Amendment was first proposed by Congress in 1789, and several states ratified shortly thereafter. However, the amendment did not collect sufficient states for ratification until a new campaign ensued two centuries later. The necessary 38 states finally ratified, and the Twenty-Seventh Amendment became effective in 1992. Ensuing universal recognition of the validity of this amendment is inconsistent with the view that Article V resolutions lapse with the passage of time.\(^4\)

Third: Recognition of the durability of Article V legislative resolutions is implied by the practice of inserting specific time limits in congressional amendment proposals and in state legislative applications. Some states have supplemented this with explicit recitals to the effect that unrescinded applications are unlimited as to time unless otherwise so providing.\(^3\)

Fourth: Formulating and applying a staleness rule consistently with the purposes of Article V would be impractical. There are no judicial or legal standards sufficient to guide a court in this regard. (Is five years too long? Too short? What about 15 years?) Leaving the question to Congress would undercut the convention procedure’s fundamental purpose as a mechanism for “bypassing Congress. During the 1960s, Senator Sam Ervin pointed out that some senators and academics wanted to disregard any applications more than two years old.\(^4\) This, of course, would destroy the process, since some state legislatures meet only biennially. Allowing Congress to fix a maximum life span on applications would fit the proverbial case of the fox guarding the hen-house.

Fifth: Rescission is a common procedure.\(^4\) Legislatures, or at least lobbyists, now monitor applications and do not assume that mere duration vitiates outdated ones. Legislatures becoming dissatisfied with applications can, and do, regularly rescind them.

For these reasons, we are justified in concluding that unrescinded applications do not lapse with the mere passage of time.

V. The Unrescinded BBA and Plenary Applications

The Article V Library, which operates a website at http://article5library.org,\(^5\) currently lists 28 states with unrescinded BBA applications.\(^4\) Yet as a matter of prudence, the Mississippi application should not be counted. It may be invalid because it improperly purports to dictate to the convention an up-or-down vote on prescribed language.\(^4\) Even if it is valid, its prescribed age, such past applications from Texas lawmakers remain alive and valid until such time as they are later formally rescinded.


49 The Article V Library reports 22 rescissions of balanced budget applications since 1988 alone. See Article V Convention Application Analysis, http://article5library.org/analyze.php. There have been, of course, other rescissions.

50 See supra note 35 for my reasons for relying on the Article V Library rather than other sources.


52 The Mississippi application, adopted in 1979, is available at http://article5library.org/gettext.php?doc=1184. It reads in part as follows:

Now Therefore, Be it resolved by the House of Representatives of the State of Mississippi, the Senate Concuring Therein. That we do hereby, pursuant to Article V of the Constitution of the United States, make application to the Congress of the United States to call a convention of the several states for the proposing of the following amendment to the Constitution of the United States: [proposed amendment language]

Modern scholarly opinion is split on whether prescribed language applications are valid; I am inclined to believe they are not, based both on Founding-Era practice and on subsequent case law. Guide, supra note 2, at 38-39. Cf. Caplans, supra note 2, at 107 (pointing out that there is no Founding-Era precedent for applications that “recite the text of an amendment and require the convention to adopt that language only.”).

Two commentaries arguing to the contrary are: Rappaport, supra note 6, and Stern, supra note 6.
language seems to render it inconsistent with the other 27. Those 27 differ in various ways, but none of them is really crucial. Pre-convention documents issued by separate states always have varied somewhat, but that has not prevented conventions from meeting successfully.\textsuperscript{53}

The Article V Library lists 16 states with unrescinded plenary applications.\textsuperscript{54} Nine of those states\textsuperscript{55} have BBA applications as well, so only 7 states have plenary applications but no BBA applications: Illinois, Kentucky, New Jersey, New York, Oregon, South Carolina, and Washington. But just as we eliminated Mississippi from the BBA list, we must scratch South Carolina from the plenary list. The operative resolution of its legislature’s 1832 resolution is as follows:

Resolved, That it is expedient that a Convention of the States be called as early as practicable, to consider and determine such questions of disputed power as have arisen between the States of this confederacy and the General Government.

Resolved, That the Governor be requested to transmit copies of this preamble and resolutions to the Governors of the several States, with a request that the same may be laid before the Legislatures of their respective States, and also to our Senator’s [sic] and Representatives in Congress, to be by them laid before Congress for consideration.\textsuperscript{56}

Although this resolution qualifies as a call for a convention of the states, it does not qualify as an Article V application. It is not addressed to Congress, and it does not call for a convention for proposing amendments. Moreover, it is not plenary. The convention subject matter is identified as “such questions of disputed power as have arisen between the States of this confederacy and the General Government.” A balanced budget amendment is not within the scope of that topic; nor are term limits nor several other subjects of modern interest. This leaves six plenary applications from states that have no BBA application outstanding, each of which is addressed below.

\textit{A. Illinois}

Illinois has two valid plenary applications extant. The first dates from 1861. Its relevant language reads:

\textit{WHEREAS,} although the people of the State of Illinois do not desire any change in our Federal constitution, yet as several of our sister States have indicated that they deem it necessary that some amendment should be made thereto; and whereas, in and by the fifth article of the constitution of the United States, provision is made for proposing amendments to that instrument, either by congress or by a convention; and whereas a desire has been expressed, in various parts of the United States, for a convention to propose amendments to the constitution; therefore,

Be it resolved by the General Assembly of the State of Illinois, That if application shall be made to Congress, by any of the States deeming themselves aggrieved, to call a convention, in accordance with the constitutional provision aforesaid, to propose amendments to the constitution of the United States, that the Legislature of the State of Illinois will and does hereby concur in making such application.

Essentially, this resolution expresses the Illinois state legislature’s decision to join other states’ applications, either in 1861 or in the future. It authorizes Congress to add Illinois to any other application lists.

The other extant Illinois application was adopted in 1903, during the campaign for direct election of Senators. Its relevant language is:

\textit{WHEREAS} by direct vote of the people of the State of Illinois at a general election held in said State on the 4th day of November, A.D. 1902, it was voted that this general assembly take the necessary steps under Article V of the Constitution of the United States to bring about the election of United States Senators by direct vote of the people; and

\textit{WHEREAS} Article V of the Constitution of the United States provides that on the application of the legislatures of two-thirds of the several States the Congress of the United States shall call a convention for proposing amendments:

Now, therefore, in obedience to the expressed will of the people as expressed at the said election, be it

Resolved by the senate (the house of representatives concurring herein), That application be, and is hereby, made to the Congress of the United States to call a convention for proposing amendments to the Constitution of the United States, as provided for in said Article V . . .

The preamble explains the motivating force for the resolution, but the operative words apply for a plenary convention. It is a basic rule of legal interpretation that when there are apparent inconsistencies between a preamble and operative words, if the operative words are clear (as they are here), they prevail. In this case, moreover, there really is no inconsistency because a legislative body may be motivated by an issue without necessarily limiting its response to that issue. Significantly, the Illinois legislature left this resolution in effect after adoption of the Seventeenth Amendment and has retained it to this day. Congress can therefore count Illinois among those states applying for a convention on any topic.

\textit{B. Kentucky}

Kentucky adopted its application in 1861. The Article V Library contains only an announcement of the application from the Senate’s presiding officer. It indicates that the application is not limited, but merely asks for a convention for proposing amendments. William Pullen’s 1951 study of the application process reproduces the actual wording:

\textit{WHEREAS} the people of some states feel themselves deeply aggrieved by the policy and measures which have been

\textsuperscript{53} See generally Founding-Era Conventions, supra note 2.

\textsuperscript{54} The Article V Library uses the misnomer “general” for plenary. See supra note 33 and accompanying text.

\textsuperscript{55} Indiana, Ohio, Texas, Colorado, Iowa, Kansas, Missouri, Nebraska, and Wisconsin.

\textsuperscript{56} This and the plenary applications discussed below are available at http://article5library.org/analyze.php?topic-General&res=1&gen=1&ylimit=0.
adopted by the people of some other states; and whereas an amendment of the Constitution of the United States is deemed indispensible necessary to secure them against similar grievances in the future: therefore—

Resolved, . . . That application to Congress to call a convention for proposing amendments to the Constitution of the United States, pursuant to the fifth article, thereof, be, and the same is hereby now made by this general assembly of Kentucky; and we hereby invite our sister States to unite with us without delay, in similar application to Congress.

* * * *

Resolved, If the convention be called in accordance with the provisions of the foregoing resolutions, the legislature of the Commonwealth of Kentucky suggests for the consideration of that convention, as a basis for settling existing difficulties, the adoption, by way of amendments to the Constitution, of the resolutions offered in the Senate of the United States by the Hon. John J. Crittenden.57

This language is plenary. It recites its motivation (resolution of present and future grievances) and adds a suggested amendment, but its operative words are unlimited. Because of the recital of future grievances, the Kentucky application, like that of Illinois, looks forward to consideration of future topics.

C. New Jersey

The 1861 New Jersey application was motivated by impending civil war, as its lengthy text makes clear. However, the operative language of the resolution applies for a plenary convention:

And be it resolved, That as the Union of these States is in imminent danger unless the remedies before suggested be speedily adopted, then, as a last resort, the State of New Jersey hereby makes application, according to the terms of the Constitution, of the Congress of the United States, to call a convention (of the States) to propose amendments to said Constitution.

As in the case of Illinois and Kentucky, New Jersey’s grant of authority to Congress has never been rescinded.

D. New York

The operative language of New York’s 1789 application seeks a convention:

[W]ith full powers to take the said Constitution into their consideration, and propose such amendments thereto, as they shall find best calculated to promote our common interests, and secure to ourselves and our latest [i.e., ultimate] posterity, the great and inalienable rights of mankind.58

This application is clearly plenary.

E. Oregon

Oregon’s 1901 application, like the 1903 application of Illinois, arose out of the campaign for direct election of Senators. The preamble recites direct election as its motivation, but the operative language is unlimited:

Whereas, under the present method of the election of United States Senators by the legislatures of the several states, protracted contests frequently result in no election at all, and in all cases interfering with needed state legislation; and

Whereas, Oregon in common with many of the other states has asked congress to adopt an amendment to the Constitution of the United States providing for the election of United States Senators by direct vote of the people, and said amendment has passed the House of Representatives on several occasions, but the Senate of the United States has continually refused to adopt said amendment; therefore be it

Resolved by the House of Representatives of the State of Oregon, the Senate concurring:

That the Congress of the United States is hereby asked, and urgently requested, to call a constitutional convention for the purpose of proposing amendments to the Constitution of the United States.

F. Washington

Two Washington State applications remain in effect, both dating from the direct election of Senators campaign. The 1901 application contains no preamble or other recitals. Aside from transmittal directions, it states merely:

That application be and the same is hereby made to the Congress of the United States of America to call a convention for proposing amendments to the constitution of the United States of America as authorized by Article V of the Constitution of the United States of America.

The 1903 application is similar, except that it recites a motivation:

Whereas the present method of electing a United States Senators is expensive and conducive of unnecessary delay in the passage of useful legislation; and

Whereas the will of the people can best be ascertained by direct vote of the people: Therefore,

Be it enacted by the legislature of the State of Washington, That application be, and the same is hereby, made to the Congress of the United States of America to call a convention for proposing amendments to the Constitution of the United States of America.


58 1 Annals of Congress 29-30 (May 5, 1789). The application was dated Feb. 5, 1789.
The language of each is plenary.

VI. Aggregating Plenary with Limited Applications

We now arrive at the issue of whether a plenary application may be aggregated with narrower applications. There are two questions here. The first is, “May applications limited to one or more subjects be aggregated with plenary applications to authorize a plenary convention?” The second is “May plenary applications be aggregated with those limited to one or more subjects to authorize a limited convention?”

The first question need not detain us, for the answer is a straightforward “no.” There is no historical precedent for such a result, and as Russell Caplan observes, “a state desiring a federal balanced budget may not, and likely does not, want the Constitution changed in any other respect.”59 Today, in fact, while there is widespread current interest in a limited convention, there is little desire for a plenary one. For Congress as the agent for the state legislatures to call a plenary convention in these circumstances would violate its fiduciary duties to legislatures seeking to limit the convention’s scope.

At initial inspection, answering the question of whether plenary applications may be aggregated toward a limited convention appears difficult because obvious precedent seems lacking. In pre-constitutional practice, states almost never issued plenary applications or calls. They almost universally specified the subjects a proposed convention was to consider, although those subjects sometimes were very broad. Hence there was no occasion when states aggregated plenary calls with more limited ones. Even the post-constitutional years have seen relatively few plenary applications. The first was issued in 1789 by New York60 and the last in 1929 by Wisconsin, and in the intervening centuries there were fewer than twenty.61 A closer look at historical practice, however, reveals some promising clues.

A. Founding-Era Practice

The Founders’ understanding of the word “application,” as we have seen, included requests for conventions (as in Article V), calls, commissions, and instructions.62 An Article V application is essentially a conditional commission and instruction: It directs Congress to call a convention on the topics listed in the application once a sufficient number of other legislatures agree, and it necessarily grants Congress authority to do so.63 Like other Founding-Era applications, commissions and instructions could be narrow, wider but still limited, or plenary. Consistently with the legal maxim, “The greater includes the lesser,”64 a commissioner with wider authority could participate fully in meetings restricted to subjects narrower than, but included within, the scope of his wider authority.

One relevant instance arose out of the convention known to history as the First Continental Congress (1774). The convention call appeared in a circular letter drafted by John Jay on behalf of the New York Committee of Correspondence. It read in part as follows:

Upon these reasons we conclude, that a Congress of Deputies from the colonies in general is of the utmost moment; that it ought to be assembled without delay, and some unanimous resolutions formed in this fatal emergency, not only respecting your [Boston’s] deplorable circumstances, but for the security of our common rights.65

This charge is very broad66—perhaps as close to a plenary call as any convention of states or colonies has come. Yet it is not quite plenary, because it focuses on Boston’s “deplorable circumstances” and “the security of our common rights” against Great Britain. It does not authorize discussion of, for example, colonial religious establishments or local business licensing. In response, several colonies sent commissioners to the First Continental Congress who enjoyed plenipotentiary authority—that is, they were empowered to discuss, and even to agree to, anything.67 The record reveals no doubt that the grant of plenipotentiary authority authorized commissioners to participate in a more limited convention.

Another illustration arose from the assembly in 1777 at Springfield, Massachusetts. The scope of the call included paper money, laws to prevent monopoly and economic oppression, interstate trade barriers, and “such other matters as particularly concern the immediate [w]elfare of the participating states, but it was restricted to matters “not repugnant to or interfering with the powers and authorities of the Continental Congress.”68 Connecticut, however, granted its commissioners plenipotentiary authority, omitting the restriction in the call.69 No one seems to have doubted the right of the Connecticut commissioners to participate in the convention despite their broader authority.

Similarly, the documents leading up to the 1780 Boston Convention show that it was targeted at immediate war needs. Yet New Hampshire empowered its commissioners with plenipotentiary authority to consult “on any other matters

59 Caplan, supra note 2, at 108.

60 See infra notes 72 & 73 and accompanying text for discussion.

61 The Article V Library lists 21 plenary (which it calls “general”) applications from 1788 to 1929. The first—Virginia’s 1788 application—probably does not qualify. Although it is very broad, it is limited to amendments proposed by the state ratifying conventions. Also listed is South Carolina’s 1832 resolution, but as explained above that was not an Article V application.

62 Supra notes 19-29 and accompanying text.

63 Cf. Caplan, supra note 2, at 97 (“The applications submitted under article V, therefore, are the descendants of the pre-1787 convention commissions.”).


66 Cf. Founding-Era Conventions, supra note 2, at 637.

67 Id. at 638.

68 Id. at 647.

that may be thought advisable for the public good,” and they participated fully.70

Even more on point are the first two Article V applications ever issued. The 1788 Virginia application petitioned Congress to call a convention “to take into their consideration the defects of this Constitution that have been suggested by the State Conventions.”71 This application was therefore limited. On the other hand, the 1789 New York application was plenary: It sought a convention “with full powers to take the said Constitution into their consideration, and propose such amendments thereto, as they shall find best calculated to promote our common interests, and secure to ourselves and our latest [i.e., ultimate] posterity, the great and inalienable rights of mankind.”72 The New York assembly surely intended its plenary application to aggregate with Virginia’s limited one, for the two applications were part of the same campaign for a second general convention.73 Moreover, the New York legislature was justified in so intending. When a state legislature applies to Congress for a limited convention, it grants Congress its authorization to call a convention on that topic. When a state legislature applies for a plenary convention, it grants Congress authority to call a convention to consider any amendments to the current Constitution. The plenary application says, in effect, “We’ll meet with commissioners from the other states any time to talk about whatever amendments the commissioners might think helpful.” Thus, Founding-Era practice supports the conclusion that a state issuing a plenary application thereby adds to the count for a more limited one.

B. Post-Constitutional Practice

Post-constitutional practice impels one to the same conclusion. The 1861 Washington Conference Convention was a close analogue of an Article V convention for proposing amendments: Virginia called it to propose amendments that might avert civil war. The call fixed the convention’s wide, but still limited, scope this way:

[T]o adjust the present unhappy controversies, in the spirit in which the Constitution was originally formed, and consistently with its principles, so as to afford the people of the slaveholding States adequate guarantees for the security of their rights . . . to consider, and if practicable, agree upon some suitable adjustment.74

Thus, the call provided that the subject was to (1) “adjust present . . . controversies,” provided that (2) the result was consistent with guaranteeing the “rights” of slaveholders.

The convention proceedings do not contain all of the commissioners’ credentials, but they do reproduce those issued by twelve states.75 At least ten of the twelve granted authority in excess of the scope of the call.76 Ohio, Indiana, Delaware, Pennsylvania, Rhode Island, and Missouri all authorized their commissioners to agree to “adjustments,” but without limiting their representatives to the call’s pro-slavery proviso. The four remaining states granted their commissioners authority to confer on anything:

- Illinois empowered its commissioners “to confer and consult with the Commissioners of other States who shall meet at Washington.”77
- New Jersey ordered its delegates “to confer with Congress and our sister states and urge upon them the importance of carrying into effect” certain additional statements of principle.78
- New York authorized its delegates to “confer” with those from other states “upon the complaints of any part of the country, and to suggest such remedies therefor as to them shall seem fit and proper.”79
- Massachusetts authorized its agents to “confer with the General Government, or with the separate States, or with any association of delegates from such States . . . ”80

These grants of broader power clearly were designed to commit the states to participating in a convention whose subject matter was contained within their broad grants of authority.

Still another illustration arises from the state legislatures’ campaign for direct election of U.S. Senators. The campaign ran from 1899 to 1913. During that period, many legislatures adopted applications limited to the single subject of a direct election amendment.81 Others passed plenary applications while reciting in preambles that their motivation was to obtain a direct election amendment. Three examples of such applications were discussed above in section V—those of Oregon (1901), Illinois (1903), and Washington State (1903). As in the case of the 1789 New York application, the legislatures apparently assumed that plenary applications could be aggregated with those limited to a single subject, since they issued plenary applications as vehicles for addressing a particular issue.

VII. Three Objections Answered

Article V provides that “The Congress . . . on the Application of the Legislatures of two thirds of the several States, shall call a Convention for proposing Amendments.” As the text indicates,

---

71 1 Annals of Congress 28 (May 5, 1789). The application was dated Nov. 14, 1788.
72 Id. at 29-30. The application was dated Feb. 5, 1789.
73 See Caplan, supra note 2, at 32-40.
75 Id. at 454-64.
76 Kentucky's credentials granted authority equal to the scope of the call. Id. at 457. Tennessee's credentials technically authorized only participation in a convention of the slaveholding states. Id. at 454-56.
77 Id. at 459.
78 Id. at 461.
79 Id. at 462.
80 Id. at 463-64.
this duty is ministerial and mandatory. Yet even ministerial duties may have some discretionary component. Accordingly, some may object to Congress exercising its discretion to call a convention. The first possible objection may be stated in this way:

When a legislature applies for a plenary convention, it is not announcing its willingness to discuss only narrower issues. Rather, it is asserting, “We’ll attend a convention, but only if all constitutional amendments may be considered.” Thus, a plenary application should not be taken as an application for a narrower subject.

The problem with this objection is a lack of precedent to support it. In all the history of conventions of states, I am unaware of any state that ever took this “all or nothing” position. Certainly no Article V application has ever expressed it. On the contrary, the 1789 plenary New York application and the plenary applications promoting direct election of Senators argue for the contrary.

A legislature certainly has the prerogative of taking an “all-or-nothing” position. In view of the lack of precedent, though, a legislature wishing to do so should express its position in clear language.

The second objection to aggregation may be summarized as follows:

Plenary resolutions should be scrutinized before aggregating them to see if their language is sufficiently inclusive to justify aggregation with BBA applications. If not sufficiently inclusive, they should be deemed a separate category. Thus, a plenary application that, like the 1861 Illinois resolution, looks to the future perhaps should be aggregated; but others should not be. Similarly, if an application recites a motivation other than desire for a BBA, such as direct election of Senators, then it should not be aggregated with BBA applications.

Congress (and, if need be, the courts) should reject this contention for several reasons. The initial reason involves the text and associated history. Article V provides that Congress shall call a convention “on the Application of the Legislatures of two thirds of the several States.” Running separate lists by subject would not be aggregated with BBA applications. If not sufficiently inclusive, they should be deemed a separate category. Thus, a plenary application a condition the legislature could have added, with some limited applications but not others would insert in the constitutional text. In this instance, however, there is no Founding-Era practice suggesting that the text should be read otherwise than in the most straightforward manner; an inferred exception should not be wider than the custom that implies it. This conclusion is reinforced by the Constitution’s use of the imperative: “Congress . . . shall call” and by the Founding-Era practice of treating applications in a forgiving manner.

Another reason for restraining Congress’s discretion as to which plenary applications to aggregate is the nature of Congress’ role in the convention process. When aggregating applications and issuing the call, Congress acts as an executive agent for the state legislatures. Because a primary purpose of the convention procedure is to check Congress, when it aggregates applications it does so in a conflict of interest situation. Fiduciary principles argue against allowing Congress to avoid a convention by interpretive logic chopping.

Still another reason for rejecting this second objection arises from the purpose of the convention procedure. The Founders inserted it as an important safeguard for constitutional government and for personal liberty—much like the Bill of Rights and other important constitutional checks. Just as the courts enforce most of the Bill of Rights rigorously through the use of “heightened scrutiny,” so Congress and the courts should apply heightened scrutiny to efforts to block a convention.

The third objection to aggregating plenary applications with limited applications may be stated this way:

Plenary applications should be aggregated with limited applications that already existed before the plenary applications, but not with future ones. A legislature issuing a plenary application may be on notice of previous limited applications. But it is unreasonable to assume a legislature intended to seek a convention on unknown future subjects.

This argument is stronger than the second because it offers less opportunity for Congress to block a convention by sophistic word-parsing. However, a rule that a plenary application aggregates with some limited applications but not others would insert in the plenary application a condition the legislature could have added, but chose not to. Such a rule would render plenary applications relevant for issues long past—and such as a convention to address state nullification—but irrelevant for constitutional crises that might arise in the future.

The third objection also suffers from the same lack of justification from text or precedent that attended the previous two objections. Indeed, the precedent of the Constitutional Convention cuts in the opposite direction. The Constitutional Convention was called by the Virginia general assembly in late 1786, not by Congress in February 1787 as is often claimed. The call recited as the subject matter a general overhaul of the political system. Over the next few months, state after state granted their commissioners authority to match the scope of the call. After seven states—a majority—had done so, the New York legislature restricted its commissioners to considering only amendments to the Articles of Confederation. Massachusetts imposed a similar limit even later in the process. Yet as far as we know, no one suggested the later narrow commissions abrogated the earlier broad ones. Even if the last seven states had adopted such

82 Supra note 30.
83 Roberts v. United States, 176 U.S. 222, 231 (1900) (holding that a duty can be ministerial even though its performance requires statutory construction by the officer charged with performing it).
84 See supra notes 71-73 & 81 and accompanying text.
85 Advocates of the Constitution relied heavily on the availability of the amendments convention process as a way of inducing the public to support the Constitution. Founding-Era Conventions, supra note 2, at 622-24.
86 Cf. the 1832 Georgia application.
88 Id.
89 For the credentials of the delegates to the 1787 convention, see 3 Records of the Federal Convention 559-86 (Max Farrand ed., 1937).
restrictions, thereby imposing them on the convention, the earlier states’ wider grants of authority (if not formally rescinded) would have continued those states’ commitment to the convention. The gathering would have been constrained to the narrower limits, it is true; but the commissioners with wider authority still would have been empowered and expected to participate to the extent of the convention’s scope.

A final point: In assessing all three of these objections, one must remember that if a legislature with a plenary application is dissatisfied with having that application aggregate toward a limited convention, it has several remedies:

• It may rescind or amend its application before the thirty-four state threshold is reached;
• It may join at the convention with the non-applying states in voting against any proposal; and
• It may join with non-applying states in refusing to ratify.90

VIII. Conclusion

When counting applications toward a convention for proposing a balanced budget amendment—or, indeed, toward a convention for proposing any other kind of amendment—Congress should add to the count any extant plenary applications. Currently, this count gives us 33 applications for a convention to propose a balanced budget amendment—only one short of the 34 needed to require Congress to call a convention.

90 Guide, supra note 2, at 58.
The Founders Interpret the Constitution: The Division of Federal and State Powers

By Robert G. Natelson

Note from the Editor:

This article surveys ratification-era statements by defenders of the proposed Constitution enumerating powers that would be reserved to the states. It argues that this evidence of the ratifiers’ intent sheds light on how the Constitution should be interpreted, and that it calls into question the constitutionality of much of what the federal government does today.

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.


About the Author:

Robert G. Natelson, a long-time Federalist Society member, is Professor of Law (ret.) at The University of Montana, Senior Fellow in Constitutional Jurisprudence at the Independence Institute, and a prominent originalist scholar. His constitutional research is often cited by Supreme Court justices and parties. See https://i2i.org/about/our-people/rob-natelson/. This essay is one of a series written for Federalist Society Review. Each item in the series includes at least some new material not previously appearing in the scholarly legal literature.

Perhaps the leading issue dividing proponents and opponents during the debates over whether to ratify the Constitution was whether the document granted, or could be construed to grant, excessive authority to the federal government. The well-known demand by opponents of ratification—the Antifederalists—for a bill of rights was merely one important aspect of the larger concern. To at least partially disarm this opposition, the Constitution’s advocates—the Federalists—agreed to adopt a bill of rights.

Less well known is that the Federalists repeatedly informed the public of specific powers the Constitution would leave entirely within the jurisdiction of the states and their citizens. This endeavor was not surprising or unusual, nor would it be today. After the sponsors of a legal proposal present it, they generally are asked to explain further its scope and meaning, and they respond by clarifying their proposal in greater detail. That is what the Federalists did after proposing the Constitution.

After a legal measure is adopted, the lawyers and judges interpreting it generally give great credibility to the replies given by the original sponsors. They are considered authoritative expositions and representations from those most familiar with the proposal, especially because those who adopted it likely relied upon them. The essays in The Federalist are a premier example of such replies in support of the proposed Constitution, which is one reason we prize them as guides to constitutional interpretation. Unfortunately, the replies made by Federalists enumerating powers reserved exclusively to the states have largely been overlooked.

Fifteen years ago, I found most of them in history’s recesses, dusted them off, and reproduced them in a mainstream law review article.1 Although they have excited some interest, they remain underutilized.

The enumeration of powers reserved exclusively to the states are scattered throughout the ratification records. They appeared in essays, letters, convention debates, and newspapers. They ranged in length from short expositions of one or two items to very long lists—although no author claimed to itemize every power the states would retain. The Federalists clearly intended these lists to induce public reliance. They were published and republished. They also evince a certain amount of coordination, for when the enumerations overlapped they remained remarkably consistent. This essay surveys those enumerations and adds some discovered since the 2003 article was published. This essay also focuses, as the 2003 article did not, on the relevant qualifications of the enumerators.

Before we proceed, one caveat is in order: When Federalist spokesmen—the enumerators—issued lists of powers reserved exclusively to the states, they necessarily were speaking of conditions within state boundaries. The corresponding limits

---


---
on federal territories and enclaves.2

I. ANONYMOUS ENUMERATORS

The authors of a few enumerations remain unidentified—although if the item appeared in a newspaper without attribution, the editor was probably sponsoring its thesis. An anonymous enumeration published by the Pennsylvania Gazette, Benjamin Franklin’s former newspaper, exemplifies the enumeration genre. It appeared in the issue of December 26, 1787,3 at the height of the ratification controversy:

The federal government neither makes, nor can without alteration make, any provision for the choice of probates of wills, land officers and surveyors, justices of the peace, county lieutenants, county commissioners, receivers of quit-rents, sheriffs, coroners, overseers of the poor, and constables; nor does it provide in any way for the important and innumerable tasks that must take place among the citizens of the same state, nor for criminal offenses, breaches of the peace, nuisances, or other objects of the state courts; nor for licensing marriages, and public houses; nor for county roads, nor for any other roads other than the great post roads; nor for poor-houses; nor incorporating religious and political societies, towns and boroughs; nor for charity schools, administrations on estates; and many other matters . . . 4

To restate the argument in modern terms: State governments will enjoy authority, to the exclusion of the central government, over wills and inheritance, real estate, local government, most areas of civil justice, criminal law, social services, schools, religious and political groups, local road construction, tavern licensing, and domestic relations.

Illustrating the Federalist interest in inducing public reliance was the republication of this item in the Massachusetts Gazette on January 8, 1788, the day before the Massachusetts ratifying convention commenced. It appeared under the headline, “READ THIS! READ THIS!” This list has significant overlap with an earlier list signed “A.B.” and published in the October 18, 1787 Hampshire Gazette. According to “A.B.,” the Constitution reserved to the states exclusive governmental authority over domestic relations and land titles; and over the criminal law. Federal jurisdiction would not extend to “murther [sic], adultery, theft, robbery, lying, perjury [or] defamation.”5

Other anonymous enumerators agreed that under the Constitution real estate would remain a state concern. In a long pamphlet surveying the new Constitution, “A Native of Virginia” wrote:

Congress . . . will have no power to restrain the press in any of the States . . . To [the state legislatures] is left the whole domestic government of the states; they may still regulate the rules of property, the rights of persons, every thing [sic] that relates to their internal police, and whatever effects [sic] neither foreign affairs nor the rights of other States.6

II. LAY ENUMERATORS

The identities of most of the enumerators are known. As detailed in Part III, the majority were lawyers of outstanding reputation. However, some non-lawyer enumerators also enjoyed high levels of credibility.

One example is James Madison (1751-1836). At the Virginia ratifying convention, he identified as outside the federal sphere (1) regulation of slaves and slavery,9 (2) “the law of descents,” and (3) anything that would “subvert the whole system of state laws.”10 Madison also was among those debunking the claim (still extant in some quarters) that the Necessary and Proper Clause11 added vast power to the federal store. Madison affirmed that its force was limited to the other enumerated powers.12

5 Mass. Gazette, Jan. 8, 1788, in 5 Documentary History, supra note 2, at 651. See id. at 652 for the editor’s report of the heading.


7 Id. See also “Harrington,” American Herald, Oct. 15, 1787, in 4 Documentary History, supra note 2, at 76, 79 (implying that states will have exclusive power over “real estates”); Plain Truth: Reply to An Officer of the Late Continental Army, Independent Gazetteer, Nov. 10, 1787 in 23 Documentary History, supra note 2, at 216 (reverting state reserved powers over the militia).

8 “A Native of Virginia,” Observations upon the Proposed Plan of Federal Government, Apr. 2, 1788, in 9 Documentary History, supra note 2, at 655, 691-92 (Merrill Jensen, John P. Kaminski & Gaspare J. Saladino eds., 1976-2017) (hereinafter Documentary History) (claiming that “Congress . . . will have no power to restrain the press in any of the States” and arguing that it will not do so in the capital district); 3 The Debates in the Several State Conventions on the Adoption of the Federal Constitution 40 (Jonathan Elliot ed., 2d ed. 1901) [hereinafter Elliot’s Debates] (quoting Edmund Pendleton at the Virginia ratifying convention as referring to limitations of federal power within state boundaries: “Can Congress legislate for the state of Virginia . . . or make a law altering the form of transferring property, or the rule of descents, in Virginia[?]); “Harrington,” American Herald, Oct. 15, 1787, in 4 Documentary History, supra, at 76, 79 (implying that states will have exclusive power over “real estates, provided they lay within their own limits, and the title thereof is contested by two of its own citizens”).

9 Id. at 620 (stating, in arguing against a bill of rights, “Can the general government exercise any power not delegated? . . . Does the Constitution say that they shall not alter the law of descents, or do those things which would subvert the whole system of the state laws?”).

10 U.S. Const. art. I, § 8, cl. 18.

11 3 Elliot’s Debates, supra note 2, at 453 (stating to the Virginia ratifying convention, “No power is given to the general government to interpose with respect to the property in slaves now held by the states”); id. at 621-22 (claiming there is “no warrant” for believing the federal government could emancipate slaves).

12 3 Elliot’s Debates, supra note 2, at 455 (“With respect to the supposed operation of what was denominated the sweeping clause, the gentleman, he said, was mistaken; for it only extended to the enumerated powers.
Another non-lawyer, Tench Coxe (1755-1824), was probably the Federalists’ most prolific enumerator of powers reserved exclusively to the states. He was a Philadelphia businessman and in 1789 was a Pennsylvania delegate in the Confederation Congress, which was to expire later that year. After the Constitution was ratified, he served as Alexander Hamilton’s assistant secretary of the treasury. Although little known today, Coxe was one of the most influential ratification-era essayists, and he devoted considerable ink to listing powers outside the federal sphere.

Coxe began promoting the Constitution almost as soon as it became public, writing four essays under the pen name “An American Citizen.” Then, on October 17, 1787, as “One of the People,” he first mentioned a substantive limitation on the new government: its lack of control over the press. His principal expositions on federal limits, however, appeared in three essays signed “A Freeman,” published initially in the Pennsylvania Gazette in late January and early February 1788. In these essays, Coxe sought to demonstrate that under the Constitution the states would continue to play significant roles. He listed three types of state powers: those the states would hold concurrently with the central government, those granted exclusively to the states by the Constitution’s express language (such as some election rules), and those reserved exclusively to the states by implication. Although Coxe frequently discussed these three types of powers together, only those reserved by implication are discussed here. According to “A Freeman,” the states’ exclusive domain encompassed the following reserved powers:

- “[L]aws for the inspection of the produce of the country;”
- “the making or regulation of roads, except post roads;”
- local government;
- governance of religion;
- criminal law, including nearly all mala in se and many mala prohibita;
- the law of inheritance and real property;
- control of land within state boundaries “exclusively of any interference of the federal government”;
- “all the innumerable disputes about property lying within their respective territories between their own citizens, such as titles and boundaries of lands, debts . . . mercantile contracts, & c. none of which can ever be cognizable by any department of the federal government”;
- a potpourri of other activities, including seminaries of learning, workhouses, poorhouses, hospitals, promotion of manufactures, and regulation of marriages.

As to other items, such as construction of public buildings, Coxe was not clear on whether state jurisdiction was exclusive or shared with the central government. Despite the length of Coxe’s lists, nearly all items on them were corroborated by other enumerators.

III. Lawyer Enumerators

Most of the enumerators were lawyers—indeed, pillars of the bar. In numbers and distinction they far outweighed the few attorneys on the Antifederalist side. (Even Patrick Henry, the most prominent Antifederalist attorney, was known for his passionate oratory rather than his legal ability.) The following lawyer-enumerators are grouped by their states, from south to north.

James Iredell of North Carolina (1751-1799) helped lead the pro-Constitution forces at his state’s ratifying convention. Iredell had served as a judge, law revision commissioner, and state

19 “A Freeman II,” supra note 18.
21 Id.
22 Id. See also “A Freeman III,” Pa. Gazette, Feb. 6, 1788, in 16 Documentary History, supra note 2, at 49, 51 (the “lordship of the soil”). This representation rings hollow to many modern inhabitants of Western states, where the federal government owns much of the land.
23 Id. at 510 (italics in original).
24 Id. Still another of Coxe’s enumerations, also recovered since the recovery of Enumerated, supra note 1, appeared in “A Pennsylvanian to the New York Convention,” Pa. Gazette, Jun. 11, 1788, in 20 Documentary History, supra note 2, at 1040, 1042. Its lists, which included criminal law and civil justice, overlapped those in his other writings.
25 “A Freeman I,” supra note 15. Coxe was unclear about whether state jurisdiction over construction of public buildings—including canals, bridges, ferries, light houses, wharves, libraries, and state office buildings—would be exclusive or concurrent with federal jurisdiction. The inclusion of state buildings in this part of the list suggests exclusivity, but if Coxe meant to say the federal government could not construct wharves and lighthouses, he was in error. Construction of such aids to navigation traditionally was part of “regulat[ing] Commerce.” U.S. Const. art. I, § 8, cl. 3; Robert G. Natelson, The Original Constitution: What It Actually Said and Meant 114 (3d ed. 2014) [hereinafter Original Constitution].
26 Biographical information on these enumerators can be gleaned from their entries in American National Biography Online, https://www.ahn.org/.
attorney general. From 1790 until his death, he was a justice of the U.S. Supreme Court. At the North Carolina convention and in other venues, Iredell emphasized that all authority not granted to the federal government was reserved exclusively to the states and the people. He itemized among the powers not granted (1) control of religion, 25 (2) punishment of crimes other than treason, offenses against the law of nations, or felonies on the high seas, 28 and (3) regulation of slavery. 29

Virginia’s Edmund Pendleton (1721-1803) was his Commonwealth’s leading lawyer—or, arguably, shared top honors with George Wythe. A former justice of the peace, during the ratification he was state chancellor and respected enough to earn the sobriquet “Virginia’s Mansfield” (a reference to England’s greatest chief justice). Like several other enumerators, Pendleton had served as a state law revision commissioner. He chaired the Virginia ratifying convention. In cooperation with Madison, Pendleton debunked the Antifederalist claim that the Constitution’s Necessary and Proper Clause was a vast source of federal power. He maintained that it added no authority not otherwise implied by the Constitution’s grants. 30 Pendleton also asserted that the federal government would have no jurisdiction over inheritances or real property. 31

John Marshall (1755-1835), later America’s greatest Chief Justice, was an up-and-coming Richmond attorney when elected to his state’s ratifying convention. On the convention floor, Marshall emphasized exclusive state control over the militia, land titles, personal property, and contract law. 32

After attending the University of Edinburgh in Scotland, Alexander White (1738-1804) received his legal education in London’s Inns of Court—both at the Inner Temple and at Gray’s Inn. During the colonial era, White served as King’s Attorney for Frederick County and then for the colony of Virginia. In 1788, he was a delegate to both Virginia’s ratifying convention and its second constitutional convention. The following year, he joined Virginia’s initial delegation in the First Federal Congress. 33 Of limits on the central government, White wrote:

> There are other things so clearly out of the power of Congress, that the bare recital of them is sufficient, I mean the “rights of conscience, or religious liberty—the rights of bearing arms for defence, or for killing game—the liberty of fowling, hunting and fishing—the right of altering the laws of descents and distribution of the effects of deceased persons and titles of lands and goods, and the regulation of contracts in the individual States.” . . . The freedom of speech and of the press, are likewise out of the jurisdiction of Congress. 34

Maryland’s Alexander Contee Hanson (1749-1806) was a state judge and, like Iredell and Pendleton, had been a law revision commissioner. (His compilation was known as “Hanson’s Laws.”) In 1789, he became state chancellor, a position he held until his death. Writing as “Aristides,” Hanson listed as outside the federal purview “regulations of property, the regulations of the penal law, the protection of the weak [i.e., social services], the promotion of useful arts [technology], the whole internal government of [the states’] respective republics.” 35

Pennsylvania’s James Wilson (1742-1798) had been educated at St. Andrews College (now University) in Scotland and at the College of Philadelphia (now the University of Pennsylvania). Wilson learned law from the master: his teacher, John Dickinson, was not only a leading Founder, 36 he was also Pennsylvania’s foremost lawyer. Wilson was a member of Congress, a particularly influential delegate to the Constitutional Convention, and ultimately a Justice on the U.S. Supreme Court. Although he personally favored a powerful central government, Wilson repeatedly represented that under the Constitution that

---

27 4 Elliot’s Debates, supra note 2, at 194 (telling the North Carolina ratifying convention that Congress will “have no authority to interfere in the establishment of any religion whatsoever”).

28 Id. at 219

29 Id. at 102.

30 Pendleton stated:

> I understand that clause as not going a single step beyond the delegated powers. What can it act upon? Some power given by this Constitution. If they should be about to pass a law in consequence of this clause, they must pursue some of the delegated powers, but can by no means depart from them, or arrogate any new powers; for the plain language of the clause is, to give them power to pass laws in order to give effect to the delegated powers.

3 Elliot’s Debates, supra note 2, at 441.

31 Id. at 40 (explaining to the Virginia ratifying convention that Congress cannot “intermeddle with the local, particular affairs of the states. Can Congress legislate for the state of Virginia. . . . or make a law altering the form of transferring property, or the rule of descents, in Virginia?”).

32 Marshall explained by means of rhetorical questions as follows:

> Has the government of the United States power to make laws on every subject? Does he [apparently Patrick Henry or George Mason] understand it so? Can they make laws affecting the mode of transferring property, or contracts, or claims, between citizens of the same state? Can they go beyond the delegated powers? If they were to make a law not warranted by any of the powers enumerated, it would be considered by the judges as an infringement of the Constitution which they are to guard. They would not consider such a law as coming under their jurisdiction. They would declare it void.

Id. at 555; see also id. at 554 (quoting Marshall as arguing a federal law authorizing trespass would be unconstitutional).


34 Alexander White, Winchester Va. Gazette, Feb. 22, 1788, in 8 Documentary History, supra note 2, at 401, 404 (internal quotation marks in original).

35 “Aristides,” Remarks on the Proposed Plan of a Federal Government, Jan. 31 - Mar. 27, 1788, in 15 Documentary History, supra note 2, at 517, 545. Presumably the exclusive state power to promote useful arts referred to methods other than patents and copyrights, over which the Constitution granted authority to Congress. Thus, the states would retain exclusive authority to promote technology through subsidies, monopolies, and the like.

government would enjoy only those powers enumerated and no others. Like Madison, Pendleton, and Hamilton, he defended the Necessary and Proper Clause as no more than a recital. He affirmed that, even without a bill of rights, the federal government had no power to abridge freedom of the press.

Robert R. Livingston of New York (1746-1813) had been an influential drafter of his state's constitution. He was state chancellor when elected to the New York ratifying convention; in his capacity as chancellor, he administered the presidential oath to George Washington the following year. Like many other enumerators, Livingston emphasized the states' exclusive "power over property." As from New York was Alexander Hamilton (1755/57-1804). Hamilton's fame as a Founder and Treasury Secretary may cause us to overlook that he was also prominent among the New York City legal elite. His legal reputation derived from a period even before he was admitted to practice: The note book he composed as a law student became the standard work for educating future New York law students. Like Wilson, Hamilton wanted the federal government to be more powerful than the Constitution allowed. After the ratification, Hamilton sought to promote that goal by spinning interpretive theories in a manner foreshadowing the efforts of today's results-oriented law professors. But while ratification was still pending, Hamilton was much more circumspect. Like Pendleton and others, he maintained that the Necessary and Proper Clause added no power to the central store. He affirmed that the following were outside federal authority: land transfers, inheritance, civil justice, criminal law, domestic relations, the press, and "agriculture and . . . other concerns of a similar nature.

Nathaniel Peaslee Sargent (1731-1791) of Massachusetts was a Harvard graduate then serving as a justice of his state's highest court, and he later became chief justice. In an exhaustive 1788 letter, Justice Sargent opined that the Constitution conveyed "[v]ery few" more powers than the Articles of Confederation. Reserved exclusively to the states were:

. . . Laws respecting criminal offenders in all cases, except Treason . . . The regulating Towns, parishes, Providing ministers, schools, looking after Poor persons, punishing Idlers, vagabonds . . . regulating Highways, bridges, fisheries, common fields &c . . . regulating inheritances, descent of estates, Partition of them, last wills and Testaments, executors, Administrators, and Guardians . . . determining all controversies between our own citizens, Rules of Legitimacy, marriage and divorce . . .

In other words, criminal law, local government, religion, education, social services, infrastructure, wills and inheritance, domestic relations, and most economic regulation were all outside the federal sphere.

An essay signed "Harrington" implied that real estate was within exclusive state jurisdiction. The author may have been a prominent Massachusetts attorney named Perez Morton (1751-1837). Morton was a Harvard graduate who would later be state attorney general and speaker of the state house of representatives.

Finally, Nathaniel Chipman of Vermont (1752-1843) presented an enumeration of his state's ratifying convention, which was held seven months after the thirteenth state, Rhode Island.

\[37\] 2 Elliot's Debates, supra note 2, at 454 (stating to the Pennsylvania ratifying convention, "where the powers are particularly enumerated. In the last case, the implied result is, that nothing more is intended to be given than what is so enumerated").

\[38\] Id. at 468 ("It is saying no more than that the powers we have already particularly given, shall be effectually carried into execution.").

\[39\] Id.

\[40\] Id. at 384 (disputing the claim at the New York ratifying convention that the Constitution would leave the states without power by explaining, "Is the power over property nothing? Is the power over life and death no power?").


\[42\] His most successful effort was the assertion, advanced in his Report on Manufactures, that the Taxation Clause, U.S. Const. art. I, §8, cl. 1, includes an open-ended power to spend for the "general Welfare." The Supreme Court adopted the theory over a century later in United States v. Butler, 297 U.S. 1 (1936).

\[43\] Alexander Hamilton, The Federalist No. 33 (stating of the clause that it is "only declaratory of a truth which would have resulted by necessary and unavoidable implication from the very act of constituting a federal government, and vesting it with certain specified powers").

\[44\] 2 Elliot's Debates, supra note 2, at 344 (listing at the New York convention "the administration of justice and the execution of the civil and criminal laws"); id. at 268 (claiming the Constitution does not permit the federal government to "new-model [i.e., reform] the internal police [i.e., public policy] of any state . . . alter, or abrogate . . . the whole of its civil and criminal institutions . . . penetrate the recesses of domestic life, and control, in all respects, the private conduct of individuals"). Cf. 2 Elliot's Debates, supra note 2, at 350 (saying at the New York ratifying convention that the states' "objects are merely civil and domestic—to support the legislative establishment, and to provide for the administration of the laws").

\[45\] Alexander Hamilton, The Federalist No. 17 ("The administration of private justice between the citizens of the same State, the supervision of agriculture and of other concerns of a similar nature . . . can never be desirable cares of a general jurisdiction"). If this is not a flat representation, I suspect it was designed to be taken for one.

\[46\] Little has been written on Sargent. See, however, his entry in the Political Graveyard, http://politicalgraveyard.com/bio/sargent.html.

\[47\] Nathaniel Peaslee Sargent to Joseph Badger, 1788 (exact date uncertain), in Documentary History, supra note 2, at 563, 567.

\[48\] Id. at 568.

\[49\] "Harrington," American Herald, Oct. 15, 1787, in 4 Documentary History, supra note 2, at 76, 79 (editor's note). See also Sarah Wentworth Morton, History of American Women, http://www.womeninhistoryblog.com/2012/08/sarah-wentworth-morton.html (discussing Perez Morton's wife). Morton would have had good reason to keep his name private; he was the subject of a scandal at the time. Id.
Documentary History, the power to dictate to newspapers necessarily communicates The fact that the Constitution denied to the federal government representation that the central government cannot license taverns representations should not be read narrowly, for their language would remain exclusively a state concern. But many of these Federalist assurances that, with few exceptions, criminal law in criminal law, for example, directly contradicts repeated grave doubts on the constitutionality of many of the federal and helped to induce them to ratify the Constitution.

In pamphlet form. All can be considered part of the ratification (other than Sargent’s) was issued on the convention floor or voters and ratifying convention delegates. Each representation (other than Sargent’s) was issued on the convention floor or disseminated in the newspapers or, as with Hanson’s work, printed in pamphlet form. All can be considered part of the ratification bargain insofar as they were heard and considered by the ratifiers and helped to induce them to ratify the Constitution.

Even construed narrowly, these representations cast grave doubts on the constitutionality of many of the federal government’s current operations. Extensive federal intervention in criminal law, for example, directly contradicts repeated Federalist assurances that, with few exceptions, criminal law would remain exclusively a state concern. But many of these representations should not be read narrowly, for their language necessarily communicates wider messages. For example, the representation that the central government cannot license taverns implies that it has no power to license hotels and restaurants. The reservation to states of the power to regulate agriculture logically extends to regulation of other fixed-location businesses. The fact that the Constitution denied to the federal government the power to dictate to newspapers necessarily communicates that it has no power to control what is said in other organs of mass communication. All of these examples suggest strongly that the enumerators considered the Constitution as reserving to the states alone the power to regulate local business. Furthermore, the assurance that Congress could not regulate slavery communicates the wider message—which some enumerators made explicit—that Congress could not regulate other kinds of property or domestic relations. Thus, when the necessary implications of these representations are taken into account, the charge of illegitimacy clouds a very large portion of modern federal activities.

Not can these representations be dismissed as the rantings of marginal figures. Other than perhaps Hamilton, whose centralizing views were extreme, none of these enumerators was out of the political mainstream. All were prominent and highly respected citizens. The attorneys among them—which is to say, most of them—adorned the legal profession in their respective states. Professionals of the quality of Edmund Pendleton, James Iredell, and Nathaniel Chipman understood the Constitution as a legal document. They knew the circumstances toward which its terms were addressed. They were vastly knowledgeable, and proficient in contemporaneous methods of documentary interpretation. Moreover, they spoke or wrote immediately after the Constitution was written and while it was still under active public consideration. Such facts render their representations far more probative of constitutional meaning than much of what is cited in legal commentary today—or, for that matter, taught in law school courses that pretend constitutional law began in 1803.

Indeed, on matters of constitutional meaning, if a twentieth century Supreme Court opinion (even a unanimous one) contradicts the likes of Nathaniel Peaslee Sargent or Alexander White, then my money is on Sargent and White. If the contest

50 When I wrote Enumerated, supra note 1, the Vermont volume of the Documentary History, supra note 2, had not been published, and thus I omitted Chipman’s enumeration. Of course, his enumeration occurred just outside the Ratification Era (i.e., from September 17, 1787, when the Constitution became public, until May 29, 1790, when Rhode Island became the thirteenth state to ratify), and therefore has less probative power than enumerations issued within that period. For a discussion of the extent to which the Vermont proceeding are probative of original understanding, see Robert G. Natelson, New Information on the Constitution’s Ratification—Part III: Vermont, https://12i.org/new-info-on-the-constitutions-ratification-part-iii-vermont/.

51 29 Documentary History, supra note 2, at 204 (speaking at the Vermont ratifying convention, Jan. 7, 1791).

52 Of course slave “property” was abolished by the Thirteenth Amendment, but that amendment did not alter the federal-state balance as to more defensible forms of property and domestic relationships.

53 Some modern federal operations seen by the Founders as outside the federal sphere are authorized by subsequent constitutional amendments, e.g., U.S. Const. amend. XIV, §5, but this is not true of most such operations.

54 Of course, this fact serves to increase the credibility of his acknowledgments of federal limitations.

55 That was the year of the Supreme Court’s decision in Marbury v. Madison, 5 U.S. 137 (1803), the traditional starting point for law school constitutional law courses. In fact, a sound understanding of the Constitution requires that one begin centuries earlier. See generally Original Constitution, supra note 25.

56 In United States v. Darby Lumber Co., 312 U.S. 100 (1941) a unanimous court held, apparently under the Necessary and Proper Clause, that Congress could use its Commerce Power to regulate the manufacturing process. In my view, this result virtually dictated the more famous decision the following year in Wickard v. Filburn, 317 U.S. 111 (1942), when the Court, once again apparently relying on the Necessary and Proper Clause, unanimously extended the Commerce Power to agriculture. Compare these holdings with the representations of Justice Sargent, Alexander White, and Alexander Hamilton to the effect that most economic regulation would be reserved to the states. On the inaccurate objection that the Founders didn’t understand these economic activities were closely related, see Robert G. Natelson, The Legal Meaning of “Commerce” in the Commerce Clause, 80 St. John’s L. Rev. 789 (2006).
is between Sargent and White and any modern law professor, I'll raise the ante. Material of such interpretive force is entitled to far more attention than it has received.

IMPEACHMENT: 
THE CONSTITUTION’S FIDUCIARY MEANING OF “HIGH . . . MISDEMEANORS”

By Robert G. Natelson

Note from the Editor:

This article explores the meaning of the phrase “high . . . Misdemeanors” in the Constitution’s Impeachment Clause. It concludes that the phrase denotes breaches of fiduciary duties.

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.


About the Author:

Robert G. Natelson, a long-time Federalist Society member, is Professor of Law (ret.) at The University of Montana, Senior Fellow in Constitutional Jurisprudence at the Independence Institute, and a prominent originalist scholar. His constitutional research is often cited by Supreme Court justices and parties. See https://i2i.org/about/our-people/rob-natelson/. This essay is one of a series written for Federalist Society Review. Each item in the series includes at least some new material not previously appearing in the scholarly legal literature.

Professor Natelson acknowledges the assistance of John Taylor Stafford, intern in Constitutional Studies at the Independence Institute, in editing and researching this article.

The Constitution provides expressly for three methods by which federal government officials can be removed from office: (1) elected officials may be defeated for re-election, (2) members of Congress may be expelled,1 and (3) judicial and executive officers may be removed on impeachment by the House of Representatives followed by trial and conviction by the Senate.2 The Constitution contains no standards governing the first two methods of removal. For the third method, however, the official must be guilty of “Treason, Bribery, or other high Crimes and Misdemeanors.”3

Modern commentators disagree over what the Founders meant4 by the term “high . . . Misdemeanors.” Some have argued the term comprehends only violations of the criminal law.5 Others, most famously then-Representative Gerald Ford, have claimed it encompasses whatever Congress decides it encompasses.6 Neither of these two views comports with the Constitution’s text. If the Founders understood “high . . . Misdemeanors” to be limited to criminal violations, they could have omitted the words entirely and ended the sentence with “Crimes.” If they understood “high . . . Misdemeanors” to grant unlimited discretion, they could have omitted the phrase “Treason, Bribery, or other high Crimes.”
Other commentators contend the actual standard lies between these two extremes. The text implies this is correct, but commentators have not had great success determining what that standard is. Their formulations have tended to center on vague terms without discernible legal content, such as “unacceptable risk” and “egregious abuse.”

Why have commentators not deduced a clearer standard? Perhaps politics has gotten in the way. Most modern commentary dates from the time of the Nixon and Clinton impeachments and seems influenced by whether or not the author wanted the incumbent president impeached and convicted. A more fundamental problem may be the methodology employed. Writers have attempted to deduce standards from charges in English and American impeachment cases decided from the fourteenth through the twentieth centuries; Professor Raoul Berger’s authoritative 1973 book on the impeachment process is the premier example of this methodology. However, most of the cases examined are not particularly probative of the Founders’ understanding. Those decided after the Constitution was ratified, of course, had no effect on their understanding. The value of early cases—their breach of fiduciary duty was often due to the fact that the goals and values driving the impeachment process changed over time. To recapture the founding generation’s understanding of “high . . . Misdemeanors,” we do best to limit ourselves to the events and literature of the eighteenth century. We should take heed of earlier proceedings only to the extent authors influential during the founding generation relied on them.

I. The Eighteenth Century British Background

In considering the thesis that “high . . . Misdemeanors” referred to fiduciary violations, we should draw no negative implications from the Constitution’s use of traditional phrasing rather than the more modern formulation “breach of fiduciary duty.” During the eighteenth century, the law of fiduciaries was still fragmented and without a uniform vocabulary. The phrase “breach of fiduciary duty” was very rare. To be sure, the law increasingly recognized a commonality underlying the fragments, but lawyers employed a variety of terms for fiduciary breaches, some specific and some more general. The most common broad term was “breach of trust.”

Despite the differences in vocabulary, eighteenth century British sources display a close connection between impeachment and violation of fiduciary duty. For example, Parliamentary articles of impeachment explicitly and repetitively described the accused’s conduct as a breach of trust. Thus, the first article in the impeachment against Warren Hastings—the century’s most spectacular proceeding of the kind—charged the defendant with acting “in direct Breach of his Duty, his Trust, and of existing treaties.” The articles of impeachment against the Earl of Stratford, the Earl of Oxford, and Lord Halifax similarly charged breach of trust.

Popular secondary legal sources justified impeachment as arising from breach of trust or in similar fiduciary terms.

---


9 Berger, supra note 5, at 71-72.


11 E.g., Firmage, supra note 10, at 683 (reciting fourteenth and fifteenth century cases); Sloan & Garr, supra note 10, at 427 (1974) (same).


18 Id., Jun. 14, 1701, available at http://www.british-history.ac.uk/lords-jnl/vol16/pp743-747#h3-0005 (setting forth the articles of impeachment of
Blackstone’s *Commentaries* begins its discussion of misprisions by observing that “THE first and principal [misprision] is the mal-administration of such high officers, as are in public trust and employment,” which was “usually punished by the method of parliamentary impeachment.” Richard Wooddeson, Blackstone’s successor in Oxford University’s Vinerian Chair, wrote that “such kind of misdeeds . . . as peculiarly injure the commonwealth by the abuse of high offices of trust, are the most proper, and have been the most usual grounds for this kind of prosecution.”

To be sure, British authors popular in the eighteenth century frequently listed grounds for impeachment in addition to “breach of trust.” This was because some of those grounds were criminal and other terms were available from fiduciary jurisprudence to describe the remainder. In fact, however, the non-criminal charges were invariably what we would think of as breaches of fiduciary duty. For example, Edward Coke’s *Institutes* (written in the seventeenth century, but the British Empire’s most used legal treatise until Blackstone’s *Commentaries* appeared in 1765) recited a posthumous list of “high Misdeemors” against Cardinal Woolsey; William Petyt’s *Jus Parliamentarium*, published in 1740, reproduced the charges against Woolsey, as did an anonymous author’s 1788 legal treatise entitled *The Law of Parliamentary Impeachments*. Today we would recognize every item on the list as a breach of fiduciary duty. Petyt also summarized the 1386 impeachment of William de la Pole; he did not enumerate every charge, but rather focused on items congruent with fiduciary law: self-dealing, neglect, misdirection of funds, and misuse of the pardon power.

John Comyns’ *Digest of the Laws of England* enumerated a series of “high crimes and misdemeanors.” The first consisted of violations of criminal law (i.e., “high crimes”), such as encouraging piracy and bribery. Here again, the non-criminal violations were all fiduciary breaches:

- acting outside authority, as by ratifying a peace not approved by the parties, using the Great Seal without permission, and issuing unlawful and irregular orders;
- self-dealing, such as purchasing royal lands for less than true value, purchasing and holding a plurality of offices, and acting for one’s “own profit only”;
- other sorts of disloyalty, such as recommending a prejudicial peace, endangering the navy, holding incompatible offices, and attempting to undermine the established religion;
- neglect, such as an ambassador failing in his duty to inform other ambassadors of decisions, and an admiral “neglect[ing] the Safeguard of the Sea”;
- other breaches of the duty of care, such as delaying court proceedings, giving false information to the king, refusing to carry out one’s duties, and failing to pursue instructions; and
- violations of the duty to account, such as “taking Money, &c. from a foreign Prince, without giving an Account for it,” and selling goods taken when an admiral “for his own use without accounting for a tenth to others.”

As these examples show, grounds for impeachment were not limited to criminal infractions. Indeed, the anonymous author of *Parliamentary Impeachments* found it necessary to caution readers that crimes, as well as other sorts of malfeasance, could be impeachable offenses. Nor, on the other hand, was mere political opposition a proper ground for impeachment. Although differences in political opinion doubtlessly motivated many impeachments, successful accusation and conviction demanded proof that the defendant had committed a crime.

---

24 The actual grounds were more extensive. Berger, *supra* note 5, at 12-13 (listing grounds). Michael de la Pole was the Earl of Suffolk. *Id.* at 12.

25 *Petty*, *supra* note 22, at 194.

---

19 4 William Blackstone, *Commentaries* *121* (emphasis added).


21 4 Coke, *Institutes*, at 89-95. For the list, see *infra* note 22. That this posthumous proceeding was thought of as the equivalent of impeachment is confirmed in Anonymous (“A Barrister at Law”), *The Law of Parliamentary Impeachments* 6 (1788) (describing this proceeding as an impeachment).

22 William Petyt, *Jus Parliamentarium: or the Antient Power, Jurisdiction, Rights, Liberties, and Privileges of the Most High Court of Parliament* 212-22 (1741) (listing these charges: obtaining legatine authority from the Pope, *id.* at 215; making treaties without the king’s knowledge, *id.*; sending out letters in the king’s name without permission, *id.*; endangering the health of the king, *id.* at 214; limiting access to the king, *id.* at 214-15; self-dealing and excessive impositions on religious institutions, *id.* at 215; sowing dissension among nobles, *id.* at 219; and “by his Cruelty, Inquity, Affection, and Partiality, ha[ving] subverted the due Course and Order of your Grace’s Laws, to the undoing of a great Number of [the king’s] loving People,” *id.* at 222) (emphasis added).

23 *Parliamentary Impeachments*, *supra* note 21, at 6-12.

24 The actual grounds were more extensive. Berger, *supra* note 5, at 12-13 (listing grounds). Michael de la Pole was the Earl of Suffolk. *Id.* at 12.

25 *Petty*, *supra* note 22, at 194.
Gouverneur Morris added that he “was now sensible of the necessity of impeachments. . . . [The President] may be bribed by a greater interest to betray his trust.” When defending the Constitution in South Carolina, Charles Cotesworth Pinckney pointed out that impeachment would be available for federal officers who “behave amiss, or betray their public trust,” and his ally Edward Rutledge made a similar statement in the same context.

Moreover, there are very many instances of members of the founding generation linking impeachment to breaches of specific fiduciary duties. Thus, at the Virginia ratifying convention, Edmund Randolph saw it as a remedy for dishonesty, disloyalty, and self-dealing. George Nicholas and James Madison referred to it as a remedy for maladministration and violating the national interest, and Patrick Henry as a response to “violation of duty.”

On the other hand, Founders made it clear that “high . . . Misdemeanors” were neither politically defined nor limited to criminal offenses. Edmund Randolph affirmed that “No

36 Id. at 68 (italics added). For analogous formulations, see 1 FARRAND, supra note 35, at 292 (quoting a Virginia Plan provision that “The Government Senators and all officers of the United States to be liable to impeachment for mal-administration and violation of the laws—disloyalty and self-dealing”).

37 4 The Debates in the Several State Conventions on the Adoption of the Federal Constitution 281 (Jonathan Elliot ed., 2d ed. 1901) [hereinafter Elliot’s Debates] (italics added).

38 Id. at 276 (reporting that Edward Rutledge said, “If the President or the senators abused their trust, they were liable to impeachment and punishment; and the fewer that were concerned in the abuse of the trust, the more certain would be the punishment.”).

39 3 Elliot’s Debates, supra note 37, at 369 (quoting Edmund Randolph connecting impeachment to dishonesty); id. at 486 (quoting him connecting impeachment to receipt of emoluments from foreign powers—i.e., disloyalty and self-dealing).

40 Id. at 17 (quoting George Nicholas connecting impeachment to “mal-administration”); id. at 506 (quoting him connecting impeachment to violating the interest of the nation); id. at 516 (quoting James Madison to the same effect).

41 Id. at 398 (quoting Patrick Henry connecting impeachment to “violation of duty”). See also id. at 500 (quoting James Madison connecting impeachment to the President calling Senators from only a few states—i.e., partiality); id. at 512 (quoting Patrick Henry connecting impeachment to actions “derogatory to the honor or interest of their country”); id. at 506 (quoting George Nicholas comparing impeachment under the Constitution to impeachment in England to the extent that officials can be impeached for entering treaties “judged to derogate from the honor and interest of the nation”);

42 Randolph, then governor of Virginia, previously had served as state attorney general and had enjoyed a very large private practice. He served at the federal convention, in which he was the principal spokesman for the Virginia Plan. Eventually, he was to be the first Attorney General of the United States and the second Secretary of State. After resigning as
man ever thought of impeaching a man for an opinion,” 43 and
the influential Federalist essayist Tench Coxe assumed that an
officer could be impeached for conduct not interdicted by the
criminal law.44

III. Conclusion

We best capture the meaning of the phrase “high . . .
Misdemeanors” when we think of it as referring to breaches of
fiduciary duty. High misdemeanors are not limited to commission
of crimes, but they do not include mere political differences. While
violations of the criminal law provide grounds for impeachment,
high misdemeanors encompass breaches of the duties of loyalty,
good faith, and care, and of the obligations to account and to
follow instructions (including the law and Constitution) when
administering one’s office.

Secretary of State, he returned to private practice. See generally John J.

43 3 Elliot’s Debates, supra note 37, at 401.

44  Tench Coxe, “An American Citizen,” reprinted in 13 The Documentary
History of the Ratification of the Constitution 431, 434 (Merrill
(stating “[i]f the nature of his offence, besides its danger to his country,
should be criminal in itself—should involve a charge of fraud, murder or
treason—he may be tried for such crime”).
Why Constitutional Lawyers Need to Know Latin

By Robert G. Natelson

Note from the Editor:

This article discusses the role of the Latin language and other classical studies in the thought and work of the Founders, and it argues that sound constitutional interpretation requires an understanding of Latin.

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.


---

In thinking in eighteenth-century English . . . a rudimentary knowledge of Latin is highly useful; after all, every educated Englishman and American knew Latin, English words were generally closer in meaning to their Latin originals than they are today, and sometimes . . . it is apparent that an author is accustomed to formulating his thoughts in Latin.

—Forrest McDonald

Forrest McDonald was arguably the twentieth century’s greatest constitutional historian. But his case for Latin competency among those seeking to understand the Constitution is, if anything, understated.

A popular half-truth is that the framers wrote the Constitution in straightforward language that everyone can understand. If by “everyone” we mean the framers’ immediate audience—the politically-involved public of their own day—the claim is largely true. If by “everyone” we mean today’s public, or even today’s law professors and judges, the statement is entirely untrue.

The reason the Constitution’s language was so readily understandable to the founding generation but is obscure to the modern American public is that we lack much of the knowledge they possessed. Involved members of the founding generation knew, or could readily learn about, then-prevailing political practices. They were broadly aware of recent developments in America and Europe, and of the historical background of those events. They were one of the most legally sophisticated generations ever, as Edmund Burke observed in a famous parliamentary speech.

---

1 Forrest McDonald, Novus Ordo Seclorum xi (1985).

2 See, e.g., Robert G. Natelson, *Founding-Era Conventions and the Meaning of the Constitution’s “Convention for Proposing Amendments,”* 65 Fla. L. Rev. 615 (2013) (describing the founding-era’s extremely common practice of interstate conventions, knowledge of which was lost to later writers interpreting the Constitution’s amendment process).

3 Burke said:

Permit me, Sir, to add another circumstance in our colonies, which contributes no mean part towards the growth and effect of this untractable spirit. I mean their education. In no country perhaps in the world is the law so general a study. The profession itself is numerous and powerful; and in most provinces it takes the lead. The greater number of the deputies sent to the congress were lawyers. But all who read, and most do read, endeavour to obtain some smattering in that science. I have been told by an eminent bookseller, that in no branch of his business, after tracts of popular devotion, were so many books as those on the law exported to the plantations. The colonists have now fallen into the way of printing them for their own use. I hear that they have sold nearly as many of Blackstone’s Commentaries in America as in England. General Gage marks out this disposition very particularly in a letter on your table. He states, that all the people in his government are lawyers, or smatterers in law . . . .


---

About the Author:

Robert G. Natelson is Professor of Law (ret.) at The University of Montana, Senior Fellow in Constitutional Jurisprudence at the Independence Institute, and a prominent originalist scholar. His constitutional research is often cited by Supreme Court justices and parties. See https://i2i.org/about-us/people/robert-natelson/. This essay is one of a series written for the Federalist Society Review. Each item in the series includes at least some new material not previously appearing in the scholarly legal literature.

Professor Natelson dedicates this article to Dr. Miriam T. Griffin (1935-2018), Oxford’s Roman historian extraordinaire, and one of the kindest people he ever knew.
Moreover, every boy (and some girls) with educational aspirations studied the Greco-Roman classics from an early age. They were imbued with classical literature, poetry, history, philosophy, fable, and myth. Central to the curriculum was the Latin language, and Latin competency also opened the doors to the scholarship of the Medieval and early-modern worlds. During the founding era, Latin was, in a very real sense, America's second language. Despite its importance for understanding our nation's founding and Constitution, none of this knowledge—of eighteenth century practices and law or of Latin and classical studies—is prevalent among the voting public now. It is also rare among the lawyers, law professors, and judges who interpret the Constitution for the rest of us.

Later essays in this series will discuss eighteenth century law and political practice as tools of constitutional interpretation. This essay focuses on why the Latin language and, to some extent, its associated classical studies are indispensable tools for understanding the Constitution. I do not argue that everyone should study Latin, but I do contend that one should acquire a reasonable competency in the language before purporting to offer learned commentary on the Constitution. Note that this essay focuses on the value of the language to constitutional interpretation; it does not enter the long-standing debate over the extent of Latin's pedagogical benefits.7

I. LATIN AND THE FOUNDER'S MODES OF THOUGHT

Knowledge of Latin is indispensable to a full understanding of the Founders' literary references and modes of thought. Consider, by way of illustration, Gouverneur Morris, a particularly influential framor and the Constitution's final draftsman. Morris may have been less dedicated to the classics than Founders such as John Dickinson, Thomas Jefferson, and James Madison, but his early education revolved around the usual classical foci. When attending King's College (now Columbia University), his two favorite subjects were mathematics and Latin.9

---


5 Richard, supra note 4, at 8 (reporting that boys frequently began classical studies at age eight).

6 Id. at 12-38 (discussing school and college curricula).


8 Richard, supra note 4, at 24-38 (discussing leading Founders' immersion in classical studies).

9 1 Jared Sparks, Life of Gouverneur Morris 5 (1832).

Morris' writings reflect his classical training. For example, a favorite phrase of his was Medio tutissimus ibis. It means, approximately, "You will be safest if you go down the middle." It was the admonition of the sun god Apollo to his natural son, as reported in the Roman poet Ovid's delightful mock-epic, the Metamorphoses.11 Morris' repetition of it communicates something of his belief that moderation is a virtue. By knowing Morris' commitment to moderation—a commitment held in common with many other Founders12—we more readily see that he and his Constitutional Convention colleagues balanced competing values rather than (as some modern commentators have suggested13) affording primacy to any single value.

Morris' most famous written production seems also to have been affected by his classical training. When drafting the final Constitution for the convention's Committee of Style and Arrangement, Morris composed the Preamble in a special way.14 Rather than adopt the pedestrian prose typical of prior constitutional documents,15 he selected a tightly organized metric...

---

10 See Letter from Gouverneur Morris to James LaCaze (Feb. 21, 1788), in 16 THE DOCUMENTARY HISTORY OF THE RATIFICATION OF THE CONSTITUTION 171 (Merrill Jensen, John P. Kaminski & Gaspare J. Saladino eds., 1976-2017) (quoting the line); The Diary and Letters of Gouverneur Morris 585 (Anne Cary Morris ed., 1888) (quoting Morris using it in his Sept. 10, 1792 entry in his daily reports to Thomas Jefferson while they were diplomats in France); Letter from Gouverneur Morris to Lewis B. Sturgis (Nov. 1, 1814), in 7 John C. Hamilton, HISTORY OF THE REPUBLIC OF THE UNITED STATES OF AMERICA, AS TRACED IN THE WRITINGS OF ALEXANDER HAMILTON AND OF HIS CONTEMPORARIES 853 (1865) (also using the line).

11 Publius Ovidius Naso ("Ovid"), Metamorphoses, Lib. ii, line 137.

12 E.g., John Dickinson, who called moderation "a virtue, and the parent of virtues." John Dickinson, An Address to the Committee of Correspondence in Barbadoes (1760), in 1 THE POLITICAL WRITINGS OF JOHN DICKINSON, ESQUIRE 125 (John Dickinson ed., 1801).


14 The Preamble reads:

We the People of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America.

U.S. CONST. pmbl.

15 For example, compare the prosaic introduction to the Articles of Confederation:

The ARTICLES of CONFEDERATION and PERPETUAL UNION Between The States Of New Hampshire, Massachusetts-bay Rhode Island and Providence Plantations, Connecticut, New York, New Jersey, Pennsylvania, Delaware, Maryland, Virginia, North Carolina, South Carolina and Georgia.

Arts. of Confederation, Preamble.
Roman civil law, which was influential in eighteenth century jurisprudence. This is obviously true of the participating public.

II. Latin and the Law of the Founders

The Constitution is a legal document, designed to operate and be construed in the context of eighteenth century jurisprudence. Latin competency is an absolute requirement for full access to that jurisprudence. This is obviously true of the Roman civil law, which was influential in eighteenth century chancery courts and written almost entirely in Latin. It was also true of native Anglo-American jurisprudence. English law had been recorded originally in a mixture of Latin and Norman French. By the eighteenth century, French had been abandoned, but Latin remained prominent. Latin was prominent not merely in canons of construction and other maxims, most of which have since been translated. Latin headings and excerpts adorned English case reporters, parliamentary journals, and other legal texts. Long and frequent passages of unrendered Latin filled English case reporters, and many of those passages have never been translated.

For example, the Case of Mixed Money, decided in 1604 by the Privy Council, is a "must-read" for anyone who wishes to capture the meaning of Constitution's Coinage Clause. The case was originally reported in French, with extensive Latin insertions, by John Davies. In 1762, however, an English translation was published, and it accordingly became available to the founding generation. But only the French was rendered into English, so a very large portion of this translation—approximately 1,200 of 8,200 words—remained in Latin. There was no need to translate the Latin for contemporaneous readers, so it wasn't done.

Even today, you cannot read large portions of the Case of Mixed Money—or lengthy passages in many other precedents important to the Founders—unless you can read Latin. For that matter, you face the same problem if you venture into many of their secondary legal texts. The Institutes of Edward Coke, for example, is replete with untranslated Latin.

III. Latin as a Key to Constitutional English

Perhaps the greatest handicap faced by modern constitutional interpreters who don't know Latin is, as Professor McDonald suggested, the risk—or rather the certainty—of misunderstanding the Constitution's English. One can easily look up translations of isolated phrases such as habeas corpus or ex post facto. But without knowing Latin, one may never detect unsuspected signals lurking in seemingly ordinary words.

This handicap can afflict even the great. Few legal historians have been more celebrated than Leonard W. Levy. Yet even Professor Levy fell into the trap of arguing that the framers intended the Senate to direct foreign policy because founding-era

For other examples, see Del. Const. (1776) and Ga. Const. (1777). In an effort to be more inspirational, some state constitution writers prefaced their product with a declaration of rights. See, e.g., Va. Const. (1776); Md. Const. (1776). The 1780 Massachusetts constitution, written primarily by John Adams, featured a rather ponderous preamble. But no prior constitution writer accomplished quite what Gouverneur Morris did.


A similar sense is communicated by the inscriptions the Founders chose for the Great Seal—Novus ordo seclorum ("new order of the ages") adopted from Virgil, Eclogae iv, line 5 ("nouus ab integro secernens nascitur ordo") and Annuit coeptis ("He [God] has approved our undertakings"). adapted from Virgil, Aeneid, Bk. ix, line 625 ("Sapienti omen potens, audacius audire coepit").

E.g., Edward Gibbon, The Decline and Fall of the Roman Empire (1776) (second sentence) ("The frontiers of that extensive monarchy were guarded by ancient renown and disciplined valour."). The portion of sentence from "guarded" to "valour" is scanned in dactylic hexameter, the standard for Greek and Latin epic. Gibbon repeated the device throughout his work. Gibbon's scanion is closer to actual poetry than Morris', but the emotional effects of the two are similar.

The most complete collection of ratification-era writings is The Documentary History of the Ratification of the Constitution, supra note 10. See also Richard, supra note 4, at 39-43 (discussing the Founders' use of pseudonyms as symbols).

The Corpus Juris Civilis, compiled under the direction of the Eastern Roman Emperor Justinian (reigned 527-65 C.E.), was the standard source for the civil law. Several years ago, my daughter Rebecca Natenel Chertudi and I edited Justinian's works for Internet use. See https://ijl.org/constitution/roman-law-sources/. A comparatively minor part of the collection was composed in Greek, but we posted Greek passages in Latin, following Theodor Mommsen's translation.

16 E.g., S.S. Peloubet, A Collection of Legal Maxims in Law and Equity (1886) (containing an extensive list of Latin maxims with accompanying English translations).

21 Thus, the December 5, 1782 entry in the Journal of the House of Commons begins with the words, "Jovis, 5º [quinto] die Decembris; Anno 23º [vicesimo tertio] Georgii IIIº [tertit] Regis." The heading means, "Thursday, on the fifth day of December, in the twenty-third year of King George the Third."

The same practice was followed in America for a time. E.g., 1 J. N.Y. Provincial Convention 3 (beginning entry for April 21, 1775, with the phrase, "Die Veneris, 10 hora, a.m."—that is, Friday, at the hour of 10 a.m.").


26 E.g., 1 Coke Institutes 233a (quoting royal grants of hunting and forest rights) & 238b-239a (quoting a Latin statute).
writings referred to the president as a foreign affairs “agent.” Professor Levy apparently was unaware that “agent” then usually carried the Latinate meaning of “one who acts” (from agere, to do or drive), rather than denoting a representative or subordinate. Even writers careful to check etymologies may make mistakes if they are unaware of the linguistic context.

The Constitution and the records of its adoption are fairly loaded with English words that may be misconstrued in this way without a background understanding of Latin. Following are four examples from the constitutional text.

A. Perfect
This word in the Preamble has caused some to wonder how a union could be made “more perfect,” since the modern definition of the word is “without flaw.” Nothing can be more perfect if it is already perfect. The answer is that during the founding era, “perfect” usually meant “complete” (from perferre, to finish). The Preamble thereby announced that new union was to be more complete than that created by the Articles of Confederation.

B. Privileges
The unamended Constitution employs variants of this word in two other locations, but its meaning in the Privileges and Immunities Clause is the one most thoroughly misunderstood.

C. Necessary
The most important appearance of this word in the unamended Constitution is in the Necessary and Proper Clause. Modern readers struggle over it, and are apt to be confused by Thomas Jefferson’s untenable claim that “necessary” means absolutely requisite. Admittedly, it requires more than Latin to know that, in eighteenth century legal documents, “necessary” was a common way of conveying incidence: To describe a power as “necessary” to a principal power was to say it was incidental to the principal, in the way that the power to “lay and collect taxes” implies the power to hire tax collectors. It is easier to understand how “necessary” can have that meaning if you know that in Latin your necessarius (if male) or necessaria (if female) is your relative or close associate.
D. Application

This word appears in the Guarantee Clause of Article IV.\textsuperscript{38} It also appears in Article V, where it describes a state legislature’s resolution to Congress for a convention for proposing amendments.\textsuperscript{39} Today when we use “application” in the sense of communication, we usually mean that a person in a perceived inferior position is asking for something from a person in a perceived superior position. Thus, we speak of an “application” for a job or for admission to a law school. This usage has induced some writers to refer to an Article V application as a “petition.”\textsuperscript{40}

In the Founders’ Latinate English, though, “application” usually did not have the inferior-to-superior connotation. It still carried connotations from the Latin \textit{applicatio} (which in turn came from the verb \textit{applicere}, to join, fasten, attach to), which usually denoted the process of connecting one thing to another—much as we now employ the phrase “apply a bandage.” In eighteenth century English, an application could be a communication from an inferior to a superior, but it might just as well run between equals or from a superior to an inferior. So when the Constitution uses the word “application,” it is not referring to a request for a favor. On the contrary, Article V’s use of the future imperative (“Congress . . . shall call”) tells us that a legislative application for an amendments convention is better understood as a conditional command.\textsuperscript{41}

IV. Conclusion

Among the many misguided public policies that sever us from our cultural roots, perhaps none has been as perniciously effective as the decision of so many schools to deny Latin study to students who want or need it. As a result, even lawyers with the best of professional credentials commonly embark on careers as constitutional commentators without the requisite tools. For one not fortunate enough to study Latin while young, rectifying the deficiency is not easily accomplished. Acquiring Latin competency requires thousands of hours of hard work, rendered harder the older one becomes.\textsuperscript{42} For those who wish

\begin{itemize}
\item \textsuperscript{38} U.S. Const. art. IV, § 4 (“The United States . . . shall protect each of them against Invasion; and on Application of the Legislature, or of the Executive . . . against domestic Violence.”).
\item \textsuperscript{39} U.S. Const. art. V (“The Congress . . . on the Application of the Legislatures of two thirds of the several States, shall call a Convention for proposing Amendments . . . .”).
\item \textsuperscript{40} E.g., Thomas M. Durbin, Amending the U.S. Constitution: by Congress or by Constitutional [sic] Convention, Congressional Research Service, May 10, 1995, at iii, 1, 11 & 39 (referring to an Article V legislative application as a “petition”); James Kenneth Rogers, The Other Way to Amend the Constitution: The Article V Constitutional [sic] Convention Amendment Process, 30 Harvard J. L. & Pub. Pol’y 1005, 1022 (2007) (same). (I have inserted \textit{sic} to reflect the fact that a convention for proposing amendments is not a “constitutional convention.”)
\item \textsuperscript{41} On the use of “application” in the convention context, see Robert G. Natelson, Counting to Two Thirds: How Close Are We to A Convention for Proposing Amendments to the Constitution? 19 Fed. Soc. Rev. 50 (2018), https://fedsoc.org/commentary/publications/counting-to-two-thirds-how-close-are-we-to-a-convention-for-proposing-amendments-to-the-constitution.
\item \textsuperscript{42} This comes from personal testimony: I began Latin studies at age 32 while actively engaged in the practice of law. The task was not an easy one, and to be accurate constitutional analysts, I see no remedy over the short term other than laboring at the task for as long as it takes. One long-term solution should be obvious: Institutions of higher learning, including law schools, should insist on at least modest Latin competency among faculty teaching constitutional subjects or offering constitutional commentary. Trying to do a job without the proper tools can be ineffective, and may promulgate or perpetuate mistakes as to what the Constitution really means.
\end{itemize}
DID THE CONSTITUTION GRANT THE FEDERAL GOVERNMENT EMINENT DOMAIN POWER?: USING EIGHTEENTH CENTURY LAW TO ANSWER CONSTITUTIONAL QUESTIONS

By Robert G. Natelson

Note from the Editor:

This article asks whether the Constitution granted eminent domain power to the federal government and concludes that it did. The article demonstrates how to use eighteenth century legal sources to understand the Constitution in its legal context and thereby make accurate judgments about how it was originally understood.

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.


About the Author:

Robert G. Natelson is Professor of Law (ret.) at The University of Montana, Senior Fellow in Constitutional Jurisprudence at the Independence Institute, and a prominent originalist scholar. His constitutional research is often cited by Supreme Court justices and parties. See https://i2i.org/about/our-people/rob-natelson/. This essay is based partly on a section of his article Founding-Era Socialism: The Original Meaning of the Constitution’s Postal Clause, 7 Brit. J. Am. Legal Studies 1 (2018). However, like all other essays in this series written for Federalist Society Review, it contains new material not previously appearing in the scholarly legal literature. Professor Natelson acknowledges the assistance of John Taylor Stafford, intern in Constitutional Studies at the Independence Institute.

I. EIGHTEENTH CENTURY LAW AND THE CONSTITUTION

Did the Constitution as originally understood grant the federal government eminent domain authority? As to federal territories and enclaves, for which the federal government received general police power,1 the answer is clearly “yes.” As to land lying within state boundaries and outside federal enclaves, the Supreme Court held in Kohl v. United States that the federal government may exercise eminent domain, but the Court’s constitutional reasoning was unsound.2 The real answer to this question lies in founding-era jurisprudence and law books that today’s constitutional interpreters consult too rarely.

That eighteenth century jurisprudence can answer questions of constitutional interpretation should be obvious. The Constitution is a legal document. A clear majority of its framers were lawyers, and many of the rest (such as James Madison) had extensive legal knowledge. Most of the Federalists who explained the Constitution to the ratifying public were lawyers.3 Several of the leading Antifederalists, including Virginia’s Patrick Henry and New York’s Robert Yates (possibly the author of the widely distributed “Brutus” essays), were likewise members of the Bar. The Constitution contains many legal terms of art,4 and the participants in the ratification debates often explained the document in explicitly legal terms.5 Just as one of my prior essays in Federalist Society Review illustrated how knowledge of the Latin language can assist in constitutional interpretation,6 this essay illustrates how eighteenth century law can do so by exploring whether the Constitution granted the power of eminent domain to the federal government.

1 U.S. Const. art. I, § 8, cl. 17:

The Congress shall have Power . . . To exercise exclusive Legislation in all Cases whatsoever, over such District (not exceeding ten Miles square) as may, by Cession of particular States, and the Acceptance of Congress, become the Seat of the Government of the United States, and to exercise like Authority over all Places purchased by the Consent of the Legislature of the State in which the Same shall be, for the Erection of Forts, Magazines, Arsenals, dock-Yards, and other needful Buildings.

Id. art. IV, § 3, cl. 2:

The Congress shall have Power to dispose of and make all needful Rules and Regulations respecting the Territory or other Property belonging to the United States . . . .

2 91 U.S. 367 (1875).


4 E.g., U.S. Const. art. I, § 8, cl. 18 (“necessary and proper”); art. I, § 9, c. 2 (“the Writ of Habeas Corpus”); id. cl. 3 (“Bill of Attainder” and “ex post facto Law”); id. cl. 4 (“Capitation”); id. c. 5 (“duty”); art. IV, § 2, cl. 1 (“Privileges and Immunities”).

5 E.g., The Federalist No. 83 (Alexander Hamilton) (discussing rules of legal interpretation); 4 The Debates in the Several State Conventions on the Adoption of the Federal Constitution 148 (Jonathan Elliot, ed., 2d ed. 1836) (reporting James Iredell as comparing the Constitution to a “great power of attorney”).

II. Enumerated and Incidental Powers in Eighteenth Century Law

The Constitution is, of course, a document of enumerated powers. The federal government enjoys only those powers listed or incidental to those on the list. Enumerated powers can also be called express or principal powers. Under eighteenth century law, if a power was incidental to an enumerated power, then it was conveyed by implication; there was no need to set it forth expressly. An incidental power often was labeled as “needful” or “necessary” for exercise of the principal, express, or enumerated power. The Constitution itself employs both “needful” and “necessary” as synonyms for incidence.

The Constitution does not explicitly grant eminent domain authority within state boundaries. Thus, the federal government may exercise it only if it is incidental (or ancillary) to one or more express powers. It is not sufficient, as the *Kohl* court maintained, that eminent domain be “inseparable from sovereignty.” Nor is it sufficient that the Fifth Amendment’s Takings Clause qualifies its exercise, for the Clause might be merely qualifying its exercise within federal territories and enclaves.

The founding-era doctrine of principal and incidental powers was a branch of the larger jurisprudence of principal and incidental interests. Contemporaneous legal sources provide rules for determining whether an unmentioned power is incidental to an enumerated, or principal, one. The most fundamental rule was that a power was incidental if the bargain or understanding of the parties—which founding-era lawyers called the “intent of the makers”—was that it be so. When the “intent of the makers” was not known, a reviewing court adopted a default rule. The approach for deriving the default rule may be described as follows. First, the interpreter asked if the claimed incidental power was of lesser value than the enumerated one. If it was not, it could not be incidental. But if it was of lesser value, then the interpreter asked whether the claimed incidental power was tied to the enumerated power either (1) by custom, or (2) by absolute or reasonable necessity. Reasonable necessity meant that the person trying to exercise the principal power would suffer “great prejudice” in that exercise if the putatively incidental power were denied to him.

For example, suppose Abigail granted to Brianna an enumerated power to mine coal from Abigail’s land, but the grant failed to mention that Brianna had the right to use the surface for that purpose. Brianna might well claim that the right to use the surface was incidental or ancillary to the right to mine. A court assessing Brianna’s claim first would determine whether her claimed right to use the surface was less “worthy”—of lesser value—than the right to mine. If so, then the court would ask if such a right of entry was customary. If it was, then it was incidental. If it was not, then the court would ask if Brianna would suffer “great prejudice” (not mere inconvenience) without it. If so, then the right was incidental, but if Brianna would not suffer “great prejudice,” it was not. However—and this is critical—a court would not consider custom or prejudice unless the “worthiness” test was satisfied. A power was never incidental to its putative principal unless it was of lesser value than the principal.

In *McCulloch v. Maryland*, Chief Justice John Marshall followed this basic analysis. Before reaching the question of whether incorporating a national bank was “necessary,” he asked whether it was of lesser importance than the principal powers (which he called “great powers”) to which it might be incidental. In other words, he asked if the claimed power was so valuable that it would have been enumerated if the ratifiers had understood the Constitution to grant it. He concluded that authority to incorporate a national bank was of lesser value than the principal powers to which it might have been incidental. Only then did he proceed to his famous discussion of the word “necessary.”

Chief Justice John Roberts followed the same rules nearly two centuries later when deciding whether the authority to require people to purchase insurance policies was incidental to any of the Constitution’s express grants. Justice Roberts concluded that requiring people to purchase insurance was a “great power” of the kind the Constitution would have enumerated had it been granted, and that it therefore could not be incidental.

III. Was Eminent Domain a Principal Power?

A. Initial Questions

Professor William Baude has pointed out that the *Kohl* Court did not follow the procedure for determining whether an interest is incidental when it upheld the exercise of eminent domain within state boundaries. Specifically, the Court never addressed the question of whether eminent domain is too

---

7 McCulloch v. Maryland, 17 U.S. 316, 405 (1819).
9 *Id.* at 70.
10 U.S. Const. art. I, § 8, cl. 17 (“needful Buildings”) & cl. 18 (“necessary and proper”). See also Latin, supra note 6 at section III.C. (“[i]n eighteenth century legal documents, “necessary” was a common way of conveying incidence.”).
11 *Kohl*, 91 U.S. at 371-72.
12 U.S. Const. amend. V (“nor shall private property be taken for public use, without just compensation”).
14 Legal Origius, supra note 8, at 60-67.
15 *Id.* at 65.
16 E.g., Giles Jacob, A NEW LAW-DICTIONARY (10th ed. 1782) (unpaginated) (“defining incident”).
17 *McCulloch*, 17 U.S. at 407-08.
18 *Id.* at 408ff.

A year before the Supreme Court’s *Sebelius* decision, this argument as applied to the Affordable Care Act was anticipated in Robert G. Natelson & David B. Kopel, *“Health Laws of Every Description”: John Marshall’s Ruling on a Federal Health Care Law*, 12 *Engage* 49, 51 (2011).
important to be an incident.20 That question is the subject of the remainder of this essay.

Of course, if eminent domain is not a principal power, there are several principal powers to which it could be incidental. At least in some circumstances, it might be reasonably necessary for construction of offices for housing government functions. In pursuing activities under the navigation component of the Commerce Clause, Congress might need to condemn land to build lighthouses and otherwise improve harbors.21 But the most obvious role for eminent domain is in furtherance of congressional authority “to establish . . . post Roads.”22 (In the Constitution, a “post Road” is an intercity or interstate highway punctuated by rest stops; “to establish” a post road means to undertake all actions necessary to bring it into operation.)23

When the Constitution was adopted, eminent domain was a customary, and often reasonably necessary, component of road construction and improvement. Statutes empowering boards of trustees to undertake road construction and improvement routinely included grants of condemnation authority.24 However, the enumeration of condemnation power in a statute granting authority to a commission does not suggest it must be enumerated in a Constitution; one expects a statute to itemize more than a constitution.25

To understand the Founders’ view of what was and wasn’t a principal power, one must examine the law of the time. Commentators who fail to do that—who apply their reasoning ability to only a few historical scraps—may become puzzled. They may then conclude the concept of principal (or “great”) powers is tautological, incoherent, or otherwise meaningless.26 But you cannot answer constitutional questions by applying your reasoning ability, be it ever so formidable, to mere historical scraps. The Constitution was written in a legal, political, economic, and social context, and that context is key to constitutional interpretation.

20 William Baude, Re-thinking the Federal Eminent Domain Power, 122 YALE L.J. 1738 (2013) Professor Baude concluded, although on thin founding-era evidence, that eminent domain was a great power. Id. at 1755-61. The focus of his article was on later evidence.

21 During the founding era, regulation of navigation and construction of related improvements was considered part of “regulating commerce.” Robert G. Natelson, The Legal Meaning of “Commerce” In the Commerce Clause, 80 ST. JOHN’S L. REV. 789, 809-10 (2006); cf. Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1 (1824) (holding that navigation was understood by the founding generation to be part of commerce).

22 U.S. CONST. art. I, § 8, cl. 3.


24 Id. at 59.

25 Cf. McCulloch, 17 U.S. at 407 (“A constitution, to contain an accurate detail of all the subdivisions of which its great powers will admit, and of all the means by which they may be carried into execution, would partake of the prolixity of a legal code, and could scarcely be embraced by the human mind.”).

26 E.g., David S. Schwartz, A Question Perpetually Arising: Implied Powers, Capable Federalism, and the Limits of Enumerationism, 59 Ariz. L. Rev. 573, 578 & 613 (2017) (claiming that the “Great Powers theory” is an “analytical failure” and the great powers argument “is mere tautology”).

B. Eighteenth Century Law Books

Fortunately, to determine whether eminent domain was a principal power, we need not examine every law book current during the eighteenth century.27 We can limit ourselves to books that classified the law by topic and ranked the topics by importance, and focus on those known to have been popular on this side of the Atlantic. Professor Herbert A. Johnson’s survey of founding-era law libraries provides a useful measure of popularity: How many of the 22 eighteenth century American law libraries he surveyed possessed each work?28

Digests or “abridgments” organized law by “titles” and further broke down titles into divisions and sub-divisions. For an abridgment to gain popularity among lawyers, its organizational scheme had to reflect the way lawyers thought. Abridgments covered both statutory and case law, but we turn first to statutory digests because eminent domain was authorized by statute and because the new federal government would legislate that way.

Giles Jacob’s The Statute Law Commonpl’ed was one of the two most popular statutory digests in eighteenth century America.29 Many items on Jacob’s list of titles30 corresponded to powers and other concepts enunciated in the Constitution: Admiralty, Ambassadors, Appeals, Bail (also appearing in the Eighth Amendment), Bankrupts, Coin, Customs (e.g., import and export duties), Debt, Excise, Felony, Habees Corpus, Highways, Militia, Piracy, Post-office, Seamen, Soldiers, Taxes, and Weights and Measures. Other titles represented subdivisions of broader constitutional subjects; for example, Fairs and Markets, Lighthouse, Merchants and Merchandize, and Trade were all aspects of the “Power . . . to regulate Commerce . . . .”31 Still other titles—most in fact—were subjects the Constitution reserved to the states: Baron and Feme (husband and wife), Devises, Gaming, Guardians, Murder, Rape, Universities, and others. During the ratification debates, many items in this last group were identified as reserved to the states by the Constitution’s advocates.32

The coincidence of Jacob’s titles with topics mentioned in the Constitution and (if represented as reserved to the states) in the constitutional debates is striking. Although we have no direct evidence of this, it is easy to imagine the framers using Jacob’s index as a checklist, marking off some items for the new federal government and designating the rest as reserved to the states.

Partly because most of Jacob’s entries were subjects the Constitution reserved to the states, the Constitution enumerated far fewer principal powers than there were titles in Jacob’s digest.


29 Giles Jacob: The Statute Law Commonpl’ed, or A General Table to the Statutes (1739). Professor Johnson found it in five of 22 libraries surveyed, tied with Wingate’s statutory abridgment. Johnson, supra note 28, at 59.

30 Jacob, Statute Law, supra note 29 (unpaginated section following p. 409).

31 U.S. CONST. art. I, § 8, cl. 3.

32 The Founders Interpret, supra note 3.
Moreover, as shown by the division of commerce into several different titles, much of Jacob's scheme was at a lower level of generality than that of the Constitution. So a title in Jacob's digest is no guarantee that the founding generation considered the subject to be concerned with a principal power. Yet if a subject was not important enough to merit a title even in Jacob's work, this is surely evidence that the subject was not a principal one.

Jacob's digest included no title for eminent domain or for its contemporaneous synonyms—compulsory acquisition, compulsory powers, condemnation, expropriation, or taking. Buried well beneath the title “Highways” was a short reference to a statute permitting justices of the peace to condemn land for widening a highway so long as “no House, Garden, &c. be pulled down or taken away” and “Satisfaction” is paid. The subordinate location implied that eminent domain was an incident to the principal power of constructing and improving highways.

Edmund Wingate's statutory digest appeared in as many American libraries as that of Giles Jacob. A review of his volume yields similar results, except that the section on “High-Ways” included no reference to compulsory purchase of lands.

Because the latest available edition of Wingate's volume was published in 1708, well before the founding era, I also examined a later statutory digest: Thomas Walter Williams' Digest of Statute Law, published the same year the Constitution was written. Many of Williams' titles corresponded to constitutional categories. As in the Jacob abridgment, commerce was divided among several titles and eminent domain was absent. Under "Highways and Turnpikes" was a reference to purchasing land when highways needed to be widened or “turned.” If this reference included purchasing from an unwilling owner (as in Jacob’s book), then it strengthens the inference that eminent domain was considered incidental to the power to construct, relocate, and widen roads.

More comprehensive digests covered case law and some statutory matter. Probably the best of these—and one of the most popular and certainly the most current—was the five volume 1786 edition of Matthew Bacon's A New Abridgment of the Law. Many of Bacon's titles also corresponded to concepts in the Constitution, although he omitted any treatment of taxation. As in Jacob's work, some titles represented units of larger constitutional categories, and there were dozens of titles that did not correspond to constitutional categories.

There was, under “Highways,” a reference to compulsory purchase of land for highway widening, similar to that in Jacob's digest. The remaining three of the four most popular general abridgments were those by Knightly D’Anvers, Charles Viner, and John Lilly. None of these featured a title devoted to eminent domain.

We next turn to treatises that focused on real property. Professor Johnson's library survey suggests that the two most popular real property treatises were John Lilly’s Practical Conveyancer and Orlando Bridgman’s Conveyances. In third place was Edward Wood’s Complete Body of Conveyancing. Eminent domain and its synonyms did not appear among the subject titles or even in the text of any of these works. Bridgman’s and Wood’s treatises mentioned “condemnation,”

42 Matthew Bacon, A New Abridgment of the Law (Dublin, 1786).
This work appeared in ten of the 22 law libraries surveyed. Johnson, supra note 28, at 59. The abridgment by Knightly D’Anvers was more widely held (by thirteen libraries), but had not been updated since 1737, id. at 17-18, and as far as I can ascertain, was never completed. My assessment that Bacon's digest was probably the best in its category is based on my own experience with such works over the last thirteen years.

43 Postal Clause, supra note 23, at 60 n. 479.

44 For example, Carriers, Fairs and Markets, Merchant and Merchandize, Obligations, and portions of Prerogative were all subsets of the Constitution’s “Power . . . to regulate Commerce . . . .”

45 Bacon, supra note 42 (unpaginated) (located on the 141st page of a PDF file in which the title page is the first page).

46 Johnson, supra note 28, at 59 sets forth the number of libraries for each of the following: Knightly D’Anvers, A General Abridgment of the Common Law (2d ed. 1725-37) (3 vols.) (held by 13 of 22 law libraries surveyed); Charles Viner, A General Abridgment of Law and Equity (1st ed. 1742-45) (24 vols.) (held by nine libraries); John Lilly, The Practical Register, or a General Abridgment of the Law (2d ed. 1745) (two vols.) (held by eight libraries).

Lists of topics in the D’Anvers and Viner works appear in unpaginated sections at the beginning of each volume. In the two Lilly volumes, the unpaginated tables of titles appear after page 882 and 880 respectively, but before the supplemental material in each volume.


49 Edward Wood, A Compleat Body of Conveyancing in Theory and Practice (3d ed., 1770) (3 vols.). Among 22 eighteenth century libraries Johnson surveyed, Lilly’s and Bridgman’s books were each held by five and Wood’s treatise by three. Johnson, supra note 28, at 61.
but only to signify forfeiture of ship cargos for legal violations and the condemnation of individuals for violating judicial writs and for other offenses. Some other contemporaneous law books also mentioned “condemnation” in the sense of forfeiture for violating the law.

Institutes or Commentaries were treatises surveying the entire scope of the law. The two most generally held in America were William Blackstone’s Commentaries and Thomas Wood’s Institute of the Laws of England. Blackstone’s Commentaries featured a short treatment of eminent domain, identifying it as a legislative prerogative and resorting to road construction as an example. Yet Blackstone (or his editor) did not think the concept worth an index entry. Blackstone’s book had an index entry for “taking,” but it referred the reader to felonious and unlawful takings, not to eminent domain. Wood’s Institute featured no relevant entry. Newer than the institutes of Blackstone and Wood was a Systematical View of the Laws of England, by Richard Wooddeson, Blackstone’s successor at Oxford. The Systematical View did not mention eminent domain.

Another group of sources was the law dictionaries. Most of these featured comprehensive entries rather than mere definitions; they were more akin to single volume encyclopedias than to modern law dictionaries. In America, the most popular law dictionary by far was A New Law-Dictionary, compiled (like the statutory digest mentioned earlier) by the prolific Giles Jacob. Jacob’s dictionary contained definitions and accompanying discussions of most of the leading nouns (or variations thereof) in the Constitution’s enumeration of congressional powers. The entries included Tax, Debt, Money, Creditor, Commerce, Naturalization, Bankrupt [sic], Coin, Counterfeits, Post, Pirates, Letters of marque, and Militia. There was no entry for eminent domain or any of its synonyms other than “taking,” the two entries for “taking” referred to felonious and unlawful taking, as in Blackstone’s index. References to eminent domain were likewise absent in other contemporaneous law dictionaries.

In sum, the classification schemes adopted by leading works of eighteenth century law inform us that, while eminent domain was recognized as a legal concept, it was not a prominent one. Rather, it was an incident to constructing and improving highways. Eminent domain did not rank with principal powers such as taxation, military affairs, commercial regulation, bankruptcy, the post office, and road construction.

C. Eighteenth Century Instruments Granting Authority

The Constitution was only one of many founding-era documents conveying legislative authority to governments and governmental agents. Indeed, to a considerable extent, the Constitution followed patterns previously established for such instruments. The pre-constitutional instruments of this kind most relevant in America were (1) colonial charters by which the British Crown empowered colony organizers, (2) commissions by which the Crown empowered colonial governors, and (3) founding-era state constitutions, by which the people of each state created new governments and granted authority to them.

English law held that subsidiary legislative authority was within an executive’s power to govern conquered and unorganized territories. Thus, colonial charters enumerated and conveyed legislative powers to governors, usually to be exercised in

50 Bridgman, supra note 48, at 39, 231 & 310 (all referring to condemned persons); 1 Wood, supra note 49, at 358 (referring to prize goods condemned by the admiralty), 416 (condemned persons), 770 & 811 (same).

51 E.g., 2 Richard Burn, The Justice of the Peace, and Parish Officer 2-197 (15th ed., 1755) (discussing enforcement of the excise laws, including condemnation) & 227-29 (condemnation of goods for violating bans on exports of certain tools and machinery); Williams, Statutes, supra note 38, at 48, 147, 151, 155, 157 & 158 (condemnation of goods for violation of customs and excise laws); cf. 2 Nicholas Covert, The Scrivener’s Guide 745, 748 (4th ed. 1724) (containing a mortgage form by which the mortgagor covenants title against unspecified “condemnations.”

52 Johnson, supra note 28, at 59 (indicating that Blackstone’s Commentaries appeared in ten of 22 libraries, and Wood’s Institute in eight).

I have not relied on Edward Coke’s Institutes, which were widely held but already well over a century old. For the record, however, eminent domain and its synonyms do not appear as a title in eighteen century editions. Edward Coke, the First Part of the Institutes of the Laws of England (13th ed. 1788) (unpaginated table near the end of the volume); Edward Coke, the Second Part of the Institutes of the Laws of England (unnumbered ed. 1797) (unpaginated tables near beginning and end of the volume); Edward Coke, the Third Part of the Institutes of the Laws of England (unnumbered ed. 1797) (same); Edward Coke, the Fourth Part of the Institutes of the Laws of England (unnumbered ed. 1797) (same).

53 1 William Blackstone, Commentaries *135; 4 id. (unpaginated index).


55 Richard Wooddeson, A Systematical View of the Laws of England (1792). This book was published soon after the Constitution was ratified, but its lectures date from 1777. Some earlier lectures were published in Richard Wooddeson, Elements of Jurisprudence (1783). The Systematical View appeared in four of the 22 libraries Johnson surveyed. Johnson, supra note 28, at 59.

56 The edition used here is Giles Jacob, A New Law-Dictionary (10th ed. 1782). Johnson, supra note 28, at 61, states that Jacob’s dictionary in one edition or another was in twelve of 22 surveyed law libraries. Next in popularity was John Cowell’s Interpreter, held by six libraries, tied with a Law-French dictionary.

57 Principally U.S. Const. art. I, § 8, although other congressional powers are scattered throughout the document.

58 Jacob, Dictionary, supra note 56 (unpaginated).

59 John Cowell, A Law Dictionary, or the Interpreter of Words and Terms (Improved, enlarged ed. 1727) (held by six libraries); William Rastell, Les Termes de la Ley (unnumbered ed. 1742) (held by four); Timothy Cunningham, A New and Complete Law-Dictionary (1783) (held by three); Thomas Blount, A Law Dictionary and Glossary (3d ed. 1717) (held by three). I also examined two dictionaries not on Professor Johnson’s list, Anonymous, The Student’s Law-Dictionary (1740) and Richard Burn & John Burn, A New Law Dictionary (1792), with similar results.


61 Campbell v. Hall, 98 Eng. Rep. 848 (K.B. 1774) (holding that the Crown may legislate for conquered territories until formally admitting English law and institutions into the territory, but not afterward). Cf. U.S. Const. art. IV, § 3, cl. 2 (“The Congress shall have Power to dispose of and make
conjunction with an elected assembly. Typically enumerated were the powers of taxation, legislation, facilitation and regulation of commerce, land disposition, and creation of courts—all principal powers listed in the Constitution. In no colonial charter was eminent domain listed separately. Yet we know that colonial governments exercised eminent domain, so it must have been implied from enumerated authority.

In 1688, the absolutist government of James II (1685–89) issued a commission to Edmund Andros as governor of the “Dominion of New England.” The Dominion consolidated not only modern New England, but also New Jersey and New York. In addition to granting executive and judicial authority to the governor, the commission granted him an expansive list of legislative powers. These included the ability to make laws, impose taxes, appropriate funds, raise military forces, create courts, dispose of land, and provide for fairs, markets, ports, and similar instrumentalities of commerce. Eminent domain was not enumerated. This cannot be because the parties were ignorant of the subject. Only five years earlier, eminent domain was not only modern New England, but also New Jersey and New York. In addition to granting executive and judicial authority to the governor, the commission granted him an expansive list of legislative powers. These included the ability to make laws, impose taxes, appropriate funds, raise military forces, create courts, dispose of land, and provide for fairs, markets, ports, and similar instrumentalities of commerce. Eminent domain was not enumerated. This cannot be because the parties were ignorant of the subject. Only five years earlier, eminent domain had been banned in New York by an instrument revoked when the Dominion was created. Thus, it is highly unlikely that the Crown intended to deny Andros authority to take land for improvements such as roads. That authority must have been implied in the enumerated grants. In the century after the British evicted James II and the colonists disposed of Andros, the commissions of colonial governors became highly standardized. They all enumerated legislative functions to be exercised in conjunction with an elective assembly. And they all left eminent domain to implication.

Finally, between 1776 and May 29, 1790, when Rhode Island ratified the Constitution, all states except Connecticut and Rhode Island adopted new constitutions. The framers of these documents typically contemplated general purpose governments, so most state constitutions granted legislative authority in bulk rather than in enumerated detail. A partial exception was the Massachusetts Constitution of 1780, drafted primarily by John Adams. It granted to the legislature (“general court”) authority to erect a judiciary, to tax, and to otherwise legislate. Eminent domain was not set forth explicitly. But it must have been implied from the principal grants, because another portion of the same document limited its exercise.

In view of this uniform drafting history, it was perfectly reasonable for the framers to decide that eminent domain need not be enumerated, and that the Constitution would grant it by implication.

---


63 E.g., Ga. Charter (1732) (“. . . full power and authority to constitute, ordain and make, such and so many by-laws, constitutions, orders and ordinances”).

64 E.g., Pa. Charter (1681) (authorizing importation, creation of fairs, markets, and “Sea-ports, harbours, . . . and . . . other places, for discharge and unladening of goods”).

65 E.g., Ga. Charter (1732) (granting power to colonial common council to convey land). Cf. U.S. Const. art. IV, § 3, cl. 2 (“The Congress shall have Power to dispose of and make all needful Rules and Regulations respecting the Territory . . . . belonging to the United States.”).

66 E.g., Ga. Charter (1732) (“to erect and constitute judicatories and courts of record, or other courts”).

67 Stoebuck, supra note 35, at 561 n.28 (citing colonial laws authorizing condemnation for roads).


69 Id.

70 The New York Charter of Liberties and Privileges (Oct. 30, 1683), in English Historical Documents, supra note 68, at 228, 230 (denying authority to dispose of land without the owner’s consent).


The legislative authority granted was very broad:

And you the said A. B. by and with the consent of our said Council and Assembly, or the major part of them respectively, shall have full power and authority to make, constitute, and ordain laws, statutes, and ordinances, for the public peace, welfare, and good government of our said province.

Id. at 155.

Governors also arguably enjoyed legislative powers, without need for assembly consent, to “constitute” as well as appoint judges. Id. at 158; to dispose of lands, id. at 162; and to establish fairs, markets, and harbors. Id. at 163.

See also English Historical Documents, supra note 68, at 195 (editor’s note) (observing that “By the eighteenth century, the commissions of royal governors had arrived at a standard pattern, and setting forth as an example the commission of New York governor George Clinton, issued Jul. 3, 1741”).

72 E.g., Del. Const., art. 5 (1776) (granting to the legislature “all other powers necessary for the legislature of a free and independent State”); Ga. Const. (1777), art. VII (granting to the legislature “power to make such laws and regulations as may be conducive to the good order and well-being of the State”). Other constitutions without detailed enumerations of legislative powers include Md. Const. (1776), N.C. Const. (1776), N.H. Const. (1784). Part II (enumerating separately from a general legislative grant, only the power to constitute courts); N.J. Const. art II (granting indefinite legislative authority); N.Y. Const. (1777), art. II (stating a general legislative grant); Pa. Const. (1776), § 2 (granting “supreme legislative power”) & § 9 (granting to the legislature, in addition to authority to regulate its own proceedings, “all other powers necessary for the legislature of a free state or commonwealth”); S.C. Const. (1776), art. VII (general grant of legislative authority to “the president and commander-in-chief, the general assembly and legislative council”); S.C. Const. (1778), art. II (vesting legislative authority in a general assembly); Va. Const. (1776) (creating a legislature without a specific grant of authority).

73 Mass. Const. (1780), Part II, ch. I, § 1, arts III & IV.

74 Id., Part I, Art. X (requiring personal or legislative consent and reasonable compensation when eminent domain was exercised).
IV. Conclusion

The constitutional theory of principal and incidental powers was part of the jurisprudential context within which the Constitution was adopted. It is also, with the assistance of eighteenth century legal sources, quite practical to apply. Eighteenth century law recognized eminent domain as a legislative power, but not as a principal one. It was merely incidental to others, such as the authority to “establish . . . post Roads.” The Constitution did, therefore, grant by implication eminent domain authority to the federal government in the exercise of its enumerated powers.
“ADVICE” IN THE CONSTITUTION’S ADVICE AND CONSENT CLAUSE: NEW EVIDENCE FROM CONTEMPORANEOUS SOURCES

By Robert G. Natelson

Note from the Editor:

This article discusses the proper interpretation of the Constitution’s Advice and Consent Clause.

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.

I. PREVIOUS INTERPRETATIONS

The Constitution provides that certain presidential decisions are made “with the Advice and Consent of the Senate.” Article II, Section 2, Clause 2 reads as follows:

[The President] shall have Power, by and with the Advice and Consent of the Senate, to make Treaties, provided two thirds of the Senators present concur; and he shall nominate, and by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the Supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law: but the Congress may by Law vest the Appointment of such inferior Officers, as they think proper, in the President alone, in the Courts of Law, or in the Heads of Departments.

Legal commentators have spilled a fair amount of ink over the meaning of “Advice and Consent.” Some, although far from all, argue that the word “Advice” refers to senatorial input before the president presents treaties or nominations to the Senate for deliberation and approval. In a 1979 article on the treaty power, Professor Arthur Bestor contended:

On the one hand, the Senate; on the other, the President—treatymaking was to be a cooperative venture from the beginning to the end of the entire process. This, the evidence shows, was the true intent of the framers.

Other commentators have agreed that the Senate has an initiating role in the treaty and nomination processes, although most claim for the Senate a role more modest than that Professor Bestor claimed for it.

This essay examines whether the constitutional word “Advice” contemplates senatorial participation before the president presents a treaty or makes a nomination and concludes that it does not.

About the Author:

Robert G. Natelson is Professor of Law (ret.) at The University of Montana, Senior Fellow in Constitutional Jurisprudence at the Independence Institute, and a prominent originalist scholar. His constitutional research is often cited by Supreme Court justices and parties. See http://il2i.org/about-us/people/rob-natelson/. This essay is one of a series written for Federalist Society Review. Each item in the series includes at least some new material not previously in the scholarly legal literature. Professor Natelson acknowledges the assistance of John Taylor Stafford, a constitutional studies intern at the Independence Institute.

1 For arguments that “advice” does not contemplate a senatorial role in advance of presidential proposals, see John McGinnis, Appointments Clause in THE HERITAGE GUIDE TO THE CONSTITUTION 271 (2d ed. 2014); John C. Eastman, The Limited Nature of the Senate’s Advice and Consent Role, 36 U.C. Davis L. Rev. 633 (2003).


3 Howard R. Sklamberg, The Meaning of “Advice and Consent:” The Senate’s Constitutional Role in Treatymaking, 18 Mich. J. Int’l L. 445 (1997); Nicole Schwartzberg, What is a “Recess?:” Recess Appointments and the Framers’ Understanding of Advice and Consent, 28 J. L. & Pol. 231, 259-62 (2013) (concluding that the Senate was to have a strong role in treaty making, without stating specifically what that role was); Laura T. Gorjanc, Comment: The Solution to the Filibuster Problem: Putting the Advice Back in Advice and Consent, 54 Case W. Res. L. Rev. 1435, 1453 (2004) (stating that the “plain meaning” of “advice” allows the Senate to prescribe criteria for nominees).

Professor Michael D. Ramsey states that many framers thought the president and Senate would administer the treaty power in an interactive way, but “What is less clear . . . is whether the Constitution actually requires this process, or whether it is only what the Framers assumed would happen.” Michael D. Ramsey, Treaty Clause, in HERITAGE GUIDE, supra note 1, at 263, 264-65.
II. Two Errors

Although commentators contending for advance senatorial participation have examined the 1787-90 constitutional debates and pre-constitutional practice, they have misinterpreted the historical record because of two methodological errors I identified in earlier essays in this series. The first is failing to take into account changes in language. The second is failing to consult the Anglo-American jurisprudence that served as the backdrop for the Supreme Law of the Land.

Today we almost invariably think of an “agent” as a person acting on behalf of another. Several commentators have assigned that modern meaning in inappropriate contexts to the Founders’ use of the word “agent.”

Hence, when expounding on the president’s role as an “agent” in foreign affairs, commentators have understood the term to mean he would serve the Senate in the way a real estate agent represents the seller of a home—merely implementing the will of his principal. In context, however, Founders were using the word in the Latinate sense of “one who acts.” They meant only that the president would be the official who acts in foreign affairs.

Similarly, nearly all modern writers have assumed the constitutional term “Advice” means “recommendations.” This has led some to conclude the Senate should be offering, and the president considering, senatorial guidelines and other prescriptions in advance of presidential action. As explained below, however, when eighteenth century documents used “Advice” as the framers did, the word meant deliberation or consideration, so the Senate, upon receiving a proposed treaty or nomination from the president, would deliberate about the proposal (Advice) and then vote on it (Consent). Failure to notice this deliberative meaning is largely a product of the second of the common errors noted above: inattention to the jurisprudence of the time.

That jurisprudence, moreover, informs us that the correct rendition of the phrase under consideration is not “the Advice and Consent of the Senate.” Rather, it is “by and with the Advice and Consent of the Senate”—or, more succinctly, “with the Advice and Consent of the Senate.” The entire phrase means “with the deliberation and approval of the Senate.”

III. “With the Advice” in General Eighteenth Century Usage

The deliberative meanings of the noun “advice” and the verb “advise” survive in modern speech only in a few phrases, such as “take under advisement.” When the Constitution was ratified, however, both recommendatory and deliberative meanings were common. Benjamin Franklin employed both in a single sentence of his autobiography when he wrote of Pennsylvania’s governor that “He would, therefore, sometimes call in a friendly way to advise with me [i.e., deliberate with me] on difficult points, and sometimes, tho’ not often, take my advice [i.e., recommendations].”

As Franklin’s words suggest, whether the recommendatory or deliberative meaning was intended could be deduced from the context. A very important contextual factor was the presence or absence of the preposition with. That preposition usually signaled the deliberative meaning. Thus, in Samuel Johnson’s famous dictionary, the second definition for the verb “advise” was “To consider; to deliberate.” The third definition for the noun “advice” was “Consultation; deliberation: with the particle with.”

Definitions in other dictionaries were less comprehensive, but point toward similar results.

To be sure, the preposition with may not have guaranteed that the bare words “advice” or “advise” were deliberative. Nor was the preposition absolutely necessary to give those words the

---

4 E.g., Bestor, supra note 2, at 73-131.
5 E.g., Sklamberg, supra note 3, at 448-49.
7 E.g., Sklamberg, supra note 3, at 455 (1997) (stating that Alexander Hamilton’s use of the word “agent” in referring to the treaty power “suggests a limited presidential role”); Bestor, supra note 2, at 109 (reaching the same conclusion based on Madison’s use of the word “agent”).
8 Why Constitutional Lawyers Need to Know Latin, supra note 6.
9 E.g., Gorjanc, supra note 3, at 1453:

Attributing the plain meaning to the words “advice” and “consent” yields the conclusions that the Constitution allows the members of the Senate to articulate to the President the characteristics that they would prefer in his judicial nominees . . . The Oxford English Dictionary defines advice as “[o]pinion given or offered as to action; counsel.”

10 Cf. the phrase by which a king vetoed an act of Parliament: Le Roy s’advisa, meaning “The King will consider it.”


13 E.g., 1 John Ash, The New and Complete Dictionary of the English Language (London, 1775) (unpaginated) (containing only the recommendatory meaning of “advice,” but defining “advice” to mean “To consult, to consider, with with: as, "He advised with his friends"). Bailey’s dictionary contained only modern definitions for “advice,” but his entry for “advise” included the deliberative meaning “to consider or weigh in mind.” Nathan Bailey, An Universal Etymological English Dictionary (Edinburgh, 1783) (unpaginated). Sheridan’s dictionary handled the words similarly. Thomas Sheridan, A Complete Dictionary of the English Language (London, 1789) (unpaginated).

14 Several state constitutions authorized the chief executive to act “with the advice” of the executive council where no consent was required. E.g., Del. Const. (1776), art. 7 (“he [the state president] may, with the advice of the privy council, lay embargoes”); Mass. Const. (1780), Part the Second, ch. II, § 1, art. VIII (providing that the governor may issue pardons “by and with the advice of council”). One might argue “advice” in that context was recommendatory, although it may have required only deliberations in which the chief executive participated.
deliberative sense. But with greatly increased the likelihood of the deliberative sense. And if it preceded “advice and consent,” then the meaning of “advice” was almost certainly deliberative. That is why the constitutional phrase usually rendered “Advice and Consent” is better rendered “with the Advice and Consent.”

IV. “WITH THE ADVICE AND CONSENT” IN EIGHTEEN CENTURY DOCUMENTS

The phrase “with the Advice and Consent” was exceedingly common in eighteenth century writings: A search for it in the Gale database Eighteenth Century Collections Online produced 3,247 documents—of which nearly half were legal documents—and that database tends to undercount. The phrase appeared in legal instruments such as grants and charters by which one party was required to obtain the “advice and consent” of the executive council. In other cases they required the governor or president to obtain only the council’s “advice.” The small size of executive councils—ranging from four to twelve members—renders it easy for a modern interpreter to imagine members of the council actively presenting recommendations to the executive in a roundtable format. This may have encouraged the belief that “advice” had a recommendatory sense.

However, other documents show that, when used in the phrase “with the advice and consent,” the word “advice” could not have referred to consensus recommendations offered in roundtable format of the kind feasible in small executive councils. In some cases in which advice and consent were required from multiple groups, members of each group could only have deliberated with other members of their own group, since the groups were far too remote or dispersed to consult together or arrive at common recommendations. And in many cases, the entities whose “advice and consent” was required or recited were far too large to reach consensus recommendations in a roundtable setting, as some modern writers assume the Senate was to do. For example, the 1681 Pennsylvania charter empowered the “proprietary” (governor) to pass laws “by and with the advice, assent, and approbation of the Freemen of the said Country.” It seems unlikely the governor signed laws only after consulting with all of Pennsylvania’s freemen. Similarly, some instruments applied

15 E.g., John Bonar, An Inquiry into the Nature of Religious Fellowship, in The Duty and Advantage of Religious Societies 88 (1783) (pledging not to infringe or dispense with rules “unless . . . the societies with which we correspond . . . shall advise or consent thereto”); 4 The Claims of the People of England (J. Stockdale, London, 1782) (“all Resolutions taken thereupon shall be signed by such of the Privy Council as shall advise and consent to the same”).

16 Restricting the search to legal documents produced 1,456 results. Eighteenth Century Collections Online searches commonly result in undercounting because damage in the old texts causes the search engine to miss words and phrases. Searches restricted to legal documents in that database miss some legal documents because they were not classified as legal by those constructing the database.

17 E.g., 2 Edward Wood, A Compleat Body of Conveyancing in Theory and Practice 397 (London, 1770) (referring to assignment of an apprentice’s indenture “by and with the Advice and Consent of . . . his said Father”); The Royal Charter of the Dublin Society 3 (1766) (“by and with the Advice and Consent of our right trusty and right well beloved Cousin and Counsellor”).

18 E.g., Semhill v Bayly, Precedents in Chancery [Ch. 1721] 562, 563 (1750) (reciting a will: “if she shall marry with the Advice and Consent of my Executors”) (This case does not appear to be in English Reports); The Petition of William Urquhart of Meldrum 14 (1761) (referring to a grant “made with the Advice and Consent of the Barons of the Exchequer”).

19 infra notes 30 & 31 and accompanying text.

20 E.g., Mass. Charter (1691) (appointment of officials with the advice and consent of the council); cf. Pa. Charter (1681) (empowering the proprietary [governor] to pass laws “by and with the advice, assent, and approbation of theFreemen of the said Country”).

21 E.g., Del. Const. (1776), art. 9 (“The president, with the advice and consent of the privy council, may embody the militia, and act as captain-general and commander-in-chief of them, and the other military force of this State”); Mass. Const. (1780), ch I, § 1, art. IV (stating that the governor may spend money “with the advice and of the council”); Md. Const. (1776), art. XXXIII (“. . . the Governor, by and with the advice and consent of the Council, may embody the militia; and, when embodied, shall alone have the direction thereof”); N.H. Const. (1784),
the phrase “with the Advice” to one of the two chambers of the British Parliament, each of which had hundreds of members.28

When applied to legislative bodies, in fact, “with the advice and consent” seems to have referred simply to the ordinary legislative process of deliberating and voting. Thus, Parliament consisted of members who, it was said, were sent “to advise, and consent, on their behalf that sent them,”29 and parliamentary statutes began with the words, “Be it enacted by King’s most excellent Majesty, by and with the Advice and Consent of the Lords spiritual and temporal, and Commons in this present Parliament assembled.”30 Colonial legislation often began with some variation of the phrase, “Be it enacted by the Government, with the advice and consent of the general assembly.”31 At the Constitutional Convention, James Wilson suggested that treaties be approved “by and with the advice and consent of” the House of Representatives, a much larger assembly than any executive council or the proposed federal Senate.32 At the North Carolina ratifying convention, James Iredell characterized the advice and consent process thus:

The President proposes such a man for such an office. The Senate has to consider upon it. If they think him improper, the President must nominate another, whose appointment ultimately again depends upon the Senate.33

In sum, when eighteenth century records refer to a measure being adopted “with Advice and Consent” of a group, those records mean the deliberation and consent characteristic of a legislative body.

V. Dealing with Problems

In the real world, of course, the executive might seek the recommendations of key members of the legislature before making formal proposals. Roger Sherman thought the Senate might advise the president and that it was “a convenient body” to do so “from the smallness of its numbers.”34 But Sherman did not issue this statement as an interpretation of the constitutional phrase “with the Advice and Consent.” Moreover, other Founders likely would have disagreed with the proposition that the Senate was of proper size to serve as a recommendatory council.35 Although the Senate would be small compared to chambers of Parliament or the lower houses of most state legislatures, as a “kitchen cabinet” it would be unwieldy: The original thirteen states would produce 26 Senators, and the impending admission of Vermont, Kentucky, and Tennessee soon would push the number above thirty. That was triple the size of the largest state executive council.

Some framers recognized that the Senate would not serve as an executive council, and they favored a real one. At the federal convention, Gouverneur Morris suggested a Council of State consisting of six members and a secretary: “The President may from time to time submit any matter to the discussion of the Council of State . . . and may require the written opinions of any one or more of the members.”36 The convention did not adopt Morris’ idea, but it did insert a presidential power to require written opinions from department heads. The latter provision, the Opinion Clause,37 is the surviving fragment of the executive council idea. The Senate’s “Advice and Consent” role is not.

It is true that, during the first session of the First Congress, President Washington came to the Senate for “advice” or “advice and consent.” He apparently was seeking some advice in the sense of recommendations,38 but to the extent he sought advice and consent in the constitutional sense, he was asking only for senatorial consideration and approval of his proposals.39

Understanding the phrase “with the Advice and Consent” to mean “with deliberation and consent” resolves some otherwise unsettled questions. It explains why, during the ratification

28 2 Wood, supra note 17, at 136 (“with the Advice of the Lords and others of his Majesty’s most Honourable Privy Council”).
29 2 Whitelockes Notes Upon the King’s Writt 67-68 (Charles Morton ed., London 1766).
30 (Italics added). For this kind of enacting language, see The Statutes at Large From the Twenty-Sixth Year of the Reign of King George the Third 3, 7, 18 & passim (London, 1789); Wood, supra note 17, at 4; The Lords Protest on a Motion to Address His Majesty 1-2 (London, 1743) (complaining of measures adopted by the Crown “without the Advice or Consent of Parliament”); Read v. Snell [Ch. 1748] 2 Atkins 642, 654, 26 Eng. Rep. 784, 790 (“[N]othing is so undoubtedly such, as that no new laws can be made to bind the whole people of this land, but by the King, with the advice and consent of both houses of parliament, and by their united authority . . . .”); Edward Wynne, Eunomus: or, Dialogues Concerning the Law and Constitution of England 129 (London, 1785) (stating that all laws are enacted “with the advice and sent of the Lords Spiritual and Temporal, and Commons”).
31 E.g., Act of the North Carolina General Assembly concerning the election of General Assembly representatives, Nov. 28, 1746, in 4 Colonial and State Records of North Carolina 1154, available at https://docsouth.unc.edu/csr/index.php/document/csr04-0358 (“And be Enacted by his Excellency Gabriel Johnston Esqr Captain General and Governour and Commander in chief in and over this Province by and with the Advice and Consent of His Majestys Council and the General Assembly of the said Province.”); The Acts of Assembly of the Province of Pennsylvania (Philadelphia 1775), passim (“by the Proprietary and Governour, by and with the Advice and Consent of the Freeman of this Province and Territories”).
33 4 The Debates in the Several State Conventions on the Adoption of the Federal Constitution 134 (Jonathan Elliot ed., 1891).
35 At the Constitutional Convention, Sherman argued the Senate should appoint judges, 2 Farrand, supra note 32, at 41 & 43 (July 18, 1787) (James Madison). Nathaniel Gorham disagreed, contending that the Senate would be “too numerous, and too little personally responsible.” Id. at 41.
36 2 Farrand, supra note 32, at 342-43 (Aug. 20, 1787) (James Madison).
37 U.S. Const. art. II, § 2, cl. 1 (“The President . . . may require the Opinion, in writing, of the principal Officer in each of the executive Departments, upon any Subject relating to the Duties of their respective Offices”).
38 E.g., 1 Sen. Exer. J. 21-22 (Aug. 22, 1789) (asking for advice about Indian policy).
39 E.g., William Maclay, Journal of William Maclay 80, 81, 122, 127 & 282 (Edgar S. Maclay ed. 1890) ("advice and consent" expressed merely in voting).
debates, advocates of the Constitution emphasized the president’s initiating role in appointments.40

It also answers a question posed by Adam J. White.41 White has pointed out that, during the Constitutional Convention, someone suggested the “advice and consent” procedure used in Massachusetts as a model for the federal level.42 The Massachusetts constitution provided for an executive council to assist and check the governor. It further provided that appointments and financial decisions were effective only with the “advice and consent” of the council,43 while other decisions were effective merely “with the advice of council.”44 The instrument also required the council to record its “advice.”45 The text of the document makes clear that when it referred to “advice” alone, the recommendatory sense was intended. But as to those actions—appointments and financial decisions—that were valid only “with the advice and consent” of the council, the council never recorded its “advice.” All that was recorded was approval of the proposal, White observes:

In each of the Council Records entries announcing the Council’s approval of the nomination, the Council used a variation of the phrase “advised and consented to” as a whole; in no case did it specify any added “advice” beyond the mere approval of the candidate, coupled with its consenting to the nomination. . . . In not a single case do the Council Records note the council advising against spending; all entries involve the allowance of spending.46

Although White explained this as deriving from a custom by which only approving advice was recorded, there is a more persuasive explanation: The purpose of the constitutional provision requiring recording of “advice” was to put council members on record as to the recommendations they offered the governor. But “advice” in the phrase “with the Advice and Consent” did not refer to recommendations at all, but to intra-council deliberation. Because no recommendations were required for appointments and financial decisions, none was recorded.47

VI. Conclusion

Eighteenth century legal documents show that “with the Advice and Consent” was a term of art meaning “with the deliberation and consent.” When used of legislative bodies, it meant the debate and voting characteristic of legislative action. When an executive’s proposal was subject to the advice and consent of a legislative assembly, no specific action was required in advance of presentation of the proposal to the assembly. Although it is often prudent for the president to consult individual Senators before submitting a nomination or a treaty, there is no constitutional requirement that he do so.

40 E.g., The Federalist No. 66 (Alexander Hamilton):

It will be the office of the President to nominate and, with the advice and consent of the Senate, to appoint. There will, of course, be no exertion of choice on the part of the Senate. They may defeat one choice of the Executive, and oblige him to make another; but they cannot themselves choose, they can only ratify or reject the choice of the President.

41 Id. No. 76 (Alexander Hamilton) (referring to “The sole and undivided responsibility of one man” in presidential appointments; 4 Elliot, supra note 33, at 134 (quoting James Iredell):

As to offices, the Senate has no other influence but a restraint on improper appointments. The President proposes such a man for such an office. The Senate has to consider upon it. If they think him improper, the President must nominate another, whose appointment ultimately again depends upon the Senate.

42 2 Farrand, supra note 32, at 41 & 44 (July 18, 1787) (James Madison) (reporting comments by Nathaniel Gorham).

43 Mass. Const. (1780), Part the Second, ch. II, § I, art. IX (advice and consent required for appointing certain officers); id. Part the Second, ch I, § I, art. IV (advice and consent required for tax warrants), id., ch. II, § I, art. XI (advice and consent required for withdrawal of money from the treasury).

44 E.g., id. Part the Second, ch. II, § I, art. V (advice required for adjourning or proroguing the legislature); id. art. X (advice required for appointment of certain military officers).

45 Id., Part the Second, Chapter II, § 3, art. V.

46 White, supra note 41, at 137–38. (Italics in original.)

47 A stronger argument for the proposition that “with the Advice and Consent” includes an advance recommendatory component is that the framers used other words (“ratify,” “approve”) for mere deliberation and approval. However, they used “ratify” only to refer to resolutions of constitutional dimension. U.S. Const., art. V (ratification of amendments) & art. VII (ratification of the Constitution). Moreover, “with the advice and consent” was an established phrase for legislative action, so “approve” seems a more sensible term for approval of bills by a single person—the president. Id. at art. I, § 7, cl. 2.
For over thirty years, federal courts have entertained lawsuits by the two major political parties and their constituents claiming a constitutional right to voting-district boundaries that allow them to translate votes into political power. From the parties’ perspectives, the potential rewards of these so-called partisan-gerrymandering claims include the possibilities of obtaining politically favorable maps outside the legislative process and of rigging the legal framework to maximize their perceived strategic advantages.

The Supreme Court has never definitively rejected these requests for judicial assistance in winning elections and controlling the government, even though they seem unsympathetic and far afield from constitutional principle. Instead, a series of fractured decisions has allowed such claims to proceed but provided no legal standard to govern them. The result has been a series of increasingly sophisticated, expensive, and at times bizarre cases rushed through the courts, seeking to persuade Justice Kennedy to codify some new social-science metric of “fairness” into the Constitution before his retirement.

But the Supreme Court’s 2018 Gill v. Whitford decision calls this peculiar history of constitutional litigation to a close. It marks Justice Kennedy’s final vote in a partisan-gerrymandering merits case, and, more importantly, it announces that the Supreme Court has finally identified the problem with a partisan-gerrymandering claim: “It is a case about group political interests, not individual rights.” Gill holds that to state a claim of individual rights—indeed, even to state an injury to establish Article III standing—a plaintiff’s allegations must be tethered to something other than “the fortunes of political parties” and “partisan preferences.”

This ruling creates a standard too onerous for any partisan-gerrymandering plaintiff to satisfy.

A partisan-gerrymandering claim necessarily identifies an injury to a party’s statewide interests, not individual rights. The individual right to vote entails only the right to cast an equal vote for a candidate in the voter’s district, a right already protected by the one-person, one-vote principle. The additional would-be right to elect the voter’s preferred candidates can only be administered to codify some new social-science metric of “fairness” into the government, even though they seem unsympathetic and far afield from constitutional principle. Instead, a series of fractured decisions has allowed such claims to proceed but provided no legal standard to govern them. The result has been a series of increasingly sophisticated, expensive, and at times bizarre cases rushed through the courts, seeking to persuade Justice Kennedy to codify some new social-science metric of “fairness” into the Constitution before his retirement.

But the Supreme Court’s 2018 Gill v. Whitford decision calls this peculiar history of constitutional litigation to a close. It marks Justice Kennedy’s final vote in a partisan-gerrymandering merits case, and, more importantly, it announces that the Supreme Court has finally identified the problem with a partisan-gerrymandering claim: “It is a case about group political interests, not individual rights.” Gill holds that to state a claim of individual rights—indeed, even to state an injury to establish Article III standing—a plaintiff’s allegations must be tethered to something other than “the fortunes of political parties” and “partisan preferences.”

This ruling creates a standard too onerous for any partisan-gerrymandering plaintiff to satisfy.

A partisan-gerrymandering claim necessarily identifies an injury to a party’s statewide interests, not individual rights. The individual right to vote entails only the right to cast an equal vote for a candidate in the voter’s district, a right already protected by the one-person, one-vote principle. The additional would-be right to elect the voter’s preferred candidates can only be administered for groups. Moreover, because redistricting is a zero-sum game where a map favoring some interests will harm others, it can only be afforded to some groups, not all. That is so, not only as between the two major parties, but also as among the innumerable smaller interest groups that comprise those parties through the compromise necessitated by the current electoral system, under which only large, nationwide parties can hope to exert meaningful political influence. It is untenable that these groups have the constitutional right to electoral success that Democratic and Republican constituents have claimed in partisan-gerrymandering litigation. Thus, forcing partisan-gerrymandering plaintiffs to identify an individualized injury distinct from statewide partisan

2 Id.
fortunes requires them to do the impossible; explain why they deserve a greater right to vote than that afforded to other citizens.

We do not yet know what legal framework courts will eventually use to resolve these claims. But, however construed, a partisan-gerrymandering claim is a theory of party rights, not individual rights, and, worse, it implicitly assumes that voters exercise their right to vote as members of parties, not as citizens. If Gill is taken at its word, no claim of that nature can succeed.

I. PARTISAN-GERRYMANDERING LITIGATION AND ITS DISCONTENTS

A. A Brief History of Gerrymandering


In fact, the practice of crafting representational units to influence which societal constituencies are represented legislatively (and to what degree) extends back at least as far as the 1295 English Parliament, which was composed of representatives of the three “great estates” of English society: “the clergy, who were represented by two archbishops and various bishops, abbots, and archdeacons; the gentry, represented by earls and barons; and the citizens, represented by elected burgesses.” Furthermore, in English politics for hundreds of years, boroughs for representation in the House of Commons were created without regard to relative size. This allowed the creation of so-called rotten boroughs, which had “remarkably few constituents.” These were created on purpose for political reasons and were sometimes bought and sold.

In the American tradition until the 1960s, the county was the typical unit of representation, so populated urban areas were relatively underrepresented as compared to rural areas. That did not change after the ratification of the Fourteenth Amendment with its guarantee of equal protection under the laws. Southern states readmitted to the Union after Reconstruction were required to ratify the Fourteenth Amendment, but most of these states at the time of readmission had population variations from the largest to smallest legislative districts that exceeded 2 to 1, including Florida (73.7 to 1), Georgia (5.7 to 1), Louisiana (2.82 to 1), South Carolina (5.2 to 1), Texas (2.19 to 1), and North Carolina (5.2 to 1). There is no historical evidence that these population deviations were viewed at the time as posing a Fourteenth Amendment problem, and norther states too “had constitutional provisions for apportionment of at least one house of their respective legislatures which wholly disregarded the spread of population.”

American history, in addition, has seen many instances of intentional manipulation of county and political-subdivision lines and voting-district lines for political advantage. For example, five decades of Virginia politics were controlled by the “Byrd Organization,” which retained power at all levels of Virginia government through ingenious gerrymanders requiring “the support of only 5 to 7 percent of the voting-age population” for the election of Byrd operatives.

“For over 174 years the Supreme Court tenaciously refused to adjudicate districting cases involving political gerrymandering and malapportionment.” But, beginning in 1962, the Court announced and enforced the one-person, one-vote rule, requiring districts of equal population to satisfy the Fourteenth Amendment’s Equal Protection Clause. This resulted in the redistricting of virtually every legislature in the nation. Since then, legislatures have been required to redraw district lines at every level of government every ten years to account for demographic changes reflected in new census data.

The equal-population rule placed a bridle on gerrymandering. Before the Court adopted this rule, a legislature could manipulate the size of representational units and thus the weight of individual votes. It was therefore possible to guarantee outsized representation to members of favored constituencies and to deny representation to members of disfavored constituencies by assigning large numbers of disfavored voters to one representative and a small number of favored voters to one representative (or to several in small groups). No modern computer program can gerrymander so effectively.

But that possibility no longer exists. Political groups intent on rigging a map in their favor must work within the equal-population constraint, leaving limited options to impact election results using redistricting. Typically, they resort to what is known as “cracking and packing.” The political party with control of a legislature uses election-results data to identify the location of voters who have voted for and against its candidates. The party then “packs” a large number of persons who voted against it at high concentrations into a small number of districts and “cracks” the remaining persons who voted against it at low concentrations in the remaining districts.

This technique is neither new nor as effective as creating rotten boroughs. The equal-population rule gives the party that controls the redistricting a choice: it can spread out its perceived voters in order to maximize the number of districts where they constitute a majority, or it can include them at higher concentrations and ensure victory in a smaller number of districts.


5 Id.

6 Id.


8 Id. at 80.

9 Id. at 80–81.

10 Greene, supra note 4, at 1044.

11 See Vieth, 541 U.S. at 274.


13 Eaton, supra note 3, at 1196.


15 See Gill, 138 S. Ct. at 1927 (describing claims of cracking and packing in 1980s litigation).
In the latter case, the party ensures itself of victory in a limited number of districts; in the former case, the party has the possibility of a significant majority in the legislature, but it risks catastrophic losses if its perceived supporters do not turn out at expected levels or if they vote for the other side, as in a “wave” election.\textsuperscript{16}

Because of these trade-offs and uncertainties, the effects of gerrymandering are limited and tend to wane over time. Legislatures alleged to have been gerrymandered out of competitive status often see a change in party control before the end of the decade—sometimes just days after the end of litigation.\textsuperscript{17}

### B. A Brief History of Gerrymandering Litigation

Nevertheless, both major parties and their constituents have claimed in their respective turns a constitutional right to “translate their votes into seats.”\textsuperscript{18} And the Supreme Court has found itself incapable of telling the Republican and Democratic parties that they have no constitutional right to win elections. The problem, instead, has repeatedly divided the justices. In a 1986 decision, \textit{Davis v. Bandemer}, the Court allowed the claims to proceed, but under a standard sufficiently grounded in constitutional principle that neither political party could ever expect to win. In a 2004 decision, \textit{Veith v. Jubelirer}, the Court allowed the claims to proceed under no standard at all. The result has been one of the most peculiar phases of constitutional litigation in American history.

1. \textit{Davis v. Bandemer}: Partisan Gerrymandering as Akin to Racial Vote Dilution

The Supreme Court’s \textit{Davis v. Bandemer}\textsuperscript{19} decision has, for our purposes, two relevant parts. First, a majority of Justices concluded that a partisan-gerrymandering claim is justiciable. They came to that conclusion because, relying on the six-factor test of \textit{Baker v. Carr},\textsuperscript{20} they found “none of the identifying characteristics of a political question . . . present.”\textsuperscript{21} The claim raised no separation-of-powers concerns, no risk of foreign or domestic disturbance, no danger that coordinate branches of the United States government would take inconsistent positions on a question of national importance, and so forth. On the element of “judicially manageable standards”—one element among the six—the Court simply stated, quoting \textit{Baker}, “[j]udicial standards under the Equal Protection Clause are well developed and familiar.”\textsuperscript{22}

From there, the Court fractured. A three-Justice plurality proceeded by analogizing the case to the Court’s racial vote-dilution precedent. It opined that partisan-gerrymandering plaintiffs must prove themselves to be similarly situated to racial vote-dilution plaintiffs. This meant proving 1) something about the plaintiff’s group—that it is “identifiable,”\textsuperscript{23} 2) something about the state actor—that it exercised “intentional discrimination,”\textsuperscript{24} and 3) something about the alleged burden on representational rights—that the group has been “denied its chance to effectively influence the political process.”\textsuperscript{25} The third element requires much more than a showing that “a particular apportionment scheme makes it more difficult for a particular group in a particular district to elect the representatives of its choice.”\textsuperscript{26}

In articulating this standard, the plurality identified several guiding principles. One was that someone “who votes for a losing candidate is usually deemed to be adequately represented by the winning candidate and to have as much opportunity to influence that candidate as other voters in the district.”\textsuperscript{27} Another was that “a failure of proportional representation alone does not constitute impermissible discrimination under the Equal Protection Clause.”\textsuperscript{28} A third was that there is no constitutional problem with “a safe district” where the plaintiff’s group “loses election after election.”\textsuperscript{29} Based on these principles, the plurality rejected the claim before it. Even though Democratic candidates for Indiana state house seats received 51.9\% of the votes cast statewide but only 43 of 100 seats, there was no cause of action because the plaintiff had not shown that Democratic Party members in Indiana were deprived of political influence.

This plurality opinion provided the narrowest grounds for the judgment and thus, under Supreme Court procedural doctrine,\textsuperscript{30} it became the controlling opinion.\textsuperscript{31} Following that

---

\textsuperscript{16} See Nolan McCarty, Keith T. Poole & Howard Rosenthal, \textit{Does Gerrymandering Cause Polarization?}, 53 Am. J. of Pol. Sci. 666 (2009) (discussing the “dummymander,” which describes “those situations when the majority spreads its voters so thin that it actually loses seats”). A similar question arises in assessing effective minority representation under the Voting Rights Act: are minority voters better served in a smaller number of districts with higher numbers of minority voters—thereby guaranteeing their ability to elect their preferred candidates—or in a larger number of districts with lower numbers of minority voters—thereby increasing the number of districts where they may have influence but not guaranteeing their ability to elect? See Georgia v. Ashcroft, 539 U.S. 461, 480 (2003), overruled by statute Pub. L. 109–246. The rule against rotten boroughs prevents racial and political groups from being able to both guarantee ability to elect and spread out influence. Efforts to benefit some groups over others by manipulating population deviations within the equal-population rule’s leeway have created controversy. See, e.g., Harris v. Arizona Indep. Redistricting Comm’n, 136 S. Ct. 1301 (2016) (dispute over alleged manipulation of district sizes to favor racial and ethnic groups); Cox v. Larios, 542 U.S. 947 (2004) (dispute over alleged manipulation of district sizes to favor political groups). Such disputes are beyond the scope of this article.

\textsuperscript{17} See Vieth, 541 U.S. at 287 n.8.

\textsuperscript{18} Whitford v. Gill, 218 F. Supp. 3d 837, 843 (W.D. Wis. 2016).


\textsuperscript{21} Bandemer, 478 U.S. at 122.

\textsuperscript{22} Id. (quoting Baker, 369 U.S. at 226).

\textsuperscript{23} Id. at 127.

\textsuperscript{24} Id. at 192.

\textsuperscript{25} Id. at 132–33.

\textsuperscript{26} Id. at 131.

\textsuperscript{27} Id. at 132.

\textsuperscript{28} Id.

\textsuperscript{29} Id.


\textsuperscript{31} See, e.g., Republican Party of N. Carolina v. Martin, 980 F.2d 943, 955 n.22 (4th Cir. 1992).
opinion, lower courts in every single case rejected partisangerrymandering claims,32 typically on the pleadings.33 This was because, no matter how badly gerrymandering marred its fortunes, no party could prove itself similarly situated with racial vote-dilution plaintiffs. As one court put it, “even the bounds of normal political exaggeration are exceeded when the Republicans of California attempt to suggest that their political role can even be spoken of in the same breath as that of the Blacks of Burke County, Georgia and Mobile, Alabama.”34

2. Vieth v. Jubelirer: Partisan Gerrymandering Litigation as a Quest for a Manageable Standard

Bandemer satisfied no one. Legal conservatives disagreed with its justiciability ruling and were disappointed that partisan-gerrymandering cases could even be entertained. Legal progressives viewed the plurality's standard as too stringent and were disappointed that no plaintiff could win. Thus, in 2004, when the Supreme Court again addressed the question, no Justice stood by Bandemer. In Vieth v. Jubelirer, another fractured Court affirmed the grant of a 12(b)(6) motion dismissing a challenge to Pennsylvania's 2001 congressional districts.

Justice Scalia wrote for himself and three other Justices who wished to revisit and overturn Bandemer's justiciability holding. Like Bandemer, Scalia's opinion began with the six-factor Baker v. Carr test.35 But, unlike Bandemer, the opinion identified only one of those prongs as being “at issue here”—whether there are “judicially discoverable and manageable standards.”36 The plurality observed that no majority in Bandemer had identified a standard, and it summarized the history of litigation under the Bandemer plurality as “[e]ighteen years of judicial effort with no result.”37 From there, the plurality walked through “possible standards” one at a time, beginning with the Bandemer plurality and continuing through the various standards proposed by dissenting opinions.38 It rejected them all as “unmanageable,” most of them simply because the black-letter principles they articulated, such as “predominant” or “sole” purpose, were vague and indeterminate. The opinion “rejected only one on the ground that it strayed unacceptably from the Constitution's meaning.”39 There being no standard to satisfy its test, the plurality contended that the cause of action should be ruled non-justiciable.

Justice Kennedy concurred in the judgment. His opinion declined to “foreclose all possibility of judicial relief if some limited and precise rationale were found to correct an established violation of the Constitution in some redistricting cases.”40 Justice Kennedy agreed with the plurality that “the shortcomings of the other standards that have been considered to date,” including the Bandemer standard, rendered them unworkable.41 But he held out the possibility that some standard might emerge both to prevent “substantial intrusion into the Nation's political life” by the courts and to vindicate individual rights where political classifications are “applied in an invidious manner.”42 As to where that line may be, Justice Kennedy observed that the Court's decisions involving “impermissible” racial classifications are of limited relevance because “political classifications” are “generally permissible.”43 The task for future plaintiffs would be to “show an otherwise permissible classification, as applied, burdens representational rights.”44 But that was not done in the case before the Court because the plaintiffs failed to show that the political classifications were “unrelated to the aims of apportionment.”45 In all of this, Justice Kennedy was insistent that he did, in fact, “resolve this case with reference to a "standard": "The Fourteenth Amendment standard governs; and there is no doubt of that."46

The Vieth decision created confusion in the federal courts as litigants attempted to articulate a “manageable” standard that would persuade Justice Kennedy to cast his vote against allegedly gerrymandered plans. As time went on and rumors of Justice Kennedy's impending retirement swelled, these efforts became more urgent and better funded than ever.

The result has been extensive and expensive partisangerrymandering litigation in Wisconsin, North Carolina, Pennsylvania, and Maryland; cases have also been filed in 2018 in Ohio and Michigan. The cases in the first three states went to trial; the Maryland case proceeded past a motion to dismiss. None has been resolved on the pleadings, at least to date. Two of the cases resulted in district-court judgments for the plaintiffs and injunctions (which the Supreme Court promptly stayed) against districting legislation. Now, the Maryland, Wisconsin, and North

32 The cases are collected in Vieth, 541 U.S. at 280 n.6. In the only case where a claim won in district court, the Fourth Circuit reversed when elections conducted just five days after judgment directly contradicted the conclusion that North Carolina's judicial elections were persistently biased against Republican Party candidates. Republican Party of N. Carolina v. Hunt, 77 F.3d 470 (4th Cir. 1996) (unreported table decision); see Vieth, 541 U.S. at 287 n.5.


35 Vieth, 541 U.S. at 277–78.

36 Id. at 277–78.

37 Id. at 279–81.

38 Id. at 281–300.


40 Vieth, 541 U.S. at 306.

41 Id. at 308 (Kennedy, J., concurring in the judgment).

42 Id. at 307.

43 Id. at 307.

44 Id. at 314.

45 Id. at 313.

46 Id. at 313–14.
Carolina cases have all gone to the Supreme Court and are back in trial court for further proceedings.

The Pennsylvania litigation is a particularly colorful example of the reigning confusion. The case, *Agre v. Wolf*, was filed on October 2, 2017, six years after the 2011 Pennsylvania congressional redistricting and a few months after a nearly identical case was filed in state court. Notwithstanding the plaintiffs’ delay and the parallel state-court litigation over the same subject matter, the Pennsylvania three-judge panel expedited the case for trial beginning December 4, 2017—two months and two days after filing. To accomplish this, the panel suspended the rules of procedure, denying the defendants the opportunity to make motions for dismissal or summary judgment and setting discovery at a breakneck speed. As a result of the improvised proceedings, there were moments in the case where trial was conducted before the court in one room and depositions were conducted simultaneously in a nearby office. The court admitted all kinds of unusual testimony at trial, including extensive testimony by the plaintiffs expressing their wish list of “fair districts” so that they, Democratic Party members, would be represented by Democratic congresspersons. Also among the admitted evidence was testimony by an individual seeking a PhD in mechanical engineering whose redistricting experience consisted of working “on a volunteer basis for at least ten hours per week for the past nine months with a group called Concerned Citizens for Democracy that is studying gerrymandering.”

She was certified as an expert witness and testified about “five rules” she invented for what she believed would be “the best possible districting outcomes.” Practically none of the evidence the court heard was necessary. In entering judgment for the defendants, one judge concluded that the claims were non-justiciable and made no factual findings. Another concluded, based on a few points of testimony, that the plaintiffs lacked standing. The third judge dissented because he thought several districts were of sufficiently odd shape to be unconstitutional on their face; testimony into motive and expert testimony, he said, were irrelevant.

If, as the Vieth plurality believed, the “legacy of” Bandemer was “one long record of puzzlement and consternation,” the legacy of Vieth has bordered on farce.

II. *Gill v. Whitford*: A Return to the Core Question of Representational Harm

The Supreme Court’s *Gill v. Whitford* decision should bring this odd era of constitutional history to a close. This is because the decision resets the focus away from Vieth’s question of manageability and towards the core question of what the right to vote means, and it does so without assuming, as Bandemer did, that political parties can prove themselves to be similarly situated to racial vote-dilution plaintiffs. *Gill* requires a showing of how partisan gerrymandering impacts individual rights and, at the same time, demonstrates why it does not.

A. The Individualized-Harm Inquiry

The case arose as a challenge to Wisconsin’s state house and senate districting plans, drawn by Republicans in 2011. Like most partisan-gerrymandering cases, the theory of the case centered on the concept of cracking and packing and its effect on statewide vote shares. The case’s unique feature was the “efficiency gap” metric, which the plaintiffs’ lawyers argued “captures in a single number all of a district plan’s cracking and packing.” This was a new development in the social science that garnered extensive media attention.

After trial, a split three-judge district court panel entered judgment against the plan. On appeal to the Supreme Court, there were three core issues. First, the plaintiffs claimed (and the district court had approved) a statewide theory of Article III standing. The Supreme Court has held in racial-gerrymandering cases that the harm, which is derived from racial stereotyping and segregation, is experienced on a district-by-district basis, meaning that a plaintiff must reside in and challenge a specific district for representation occurs across the state. Second, there remained the unresolved question of justiciability, which meant the usual

---

50 Constitutional challenges to statewide redistricting plans are heard by three-judge panels in federal district court, 28 U.S.C. § 2284(a), with a direct appeal to the Supreme Court, 28 U.S.C. § 1253.
51 See *Agre*, 284 F. Supp. 3d at 594.
52 *Id.* at 651–57 (summarizing the plaintiffs’ testimony).
53 *Id.* at 660 (Baylson, J., dissenting).
54 *Id.* at 665.
55 *Id.* at 594 (Smith, J.).
56 *Id.* at 639 (Schwartz, J.).
57 *Id.* at 719 (Baylson, J., dissenting). The Supreme Court had summarily affirmed a case rejecting the exact same theory under Bandemer.
59 Vieth, 541 U.S. at 282.
60 *Gill*, 138 S. Ct. at 854.
Vieth arguments and counter-arguments about manageable standards. Third, there was the question of a standard and proof. The lower court had merged several equal-protection and free-speech theories together to find the following elements sufficient to prove a claim: (1) intent to crack and pack, (2) discriminatory effect in the form of a lasting majority for the party that engaged in gerrymandering, and (3) no neutral explanation.63 The lower court relied heavily on the efficiency gap and cracking and packing theories in finding liability on these elements.

The Supreme Court unanimously agreed with the state on the standing issue and therefore vacated the lower court’s judgment. The opinion by Chief Justice John Roberts approved the analogy to the Court’s racial-gerrymandering cases and held that the plaintiffs’ claim that “their votes have been diluted” alleges a harm that “arises from the particular composition of the voter’s own district, which causes his vote—having been packed or cracked—to carry less weight than it would carry in another, hypothetical district.”64

The Court then proceeded to address the plaintiffs’ contention that “their legal injury is not limited to the injury that they have suffered as individual voters, but extends also to the statewide harm to their interest ‘in their collective representation in the legislature,’ and in influencing the legislature’s overall ‘composition and policymaking.’”65 The problem with this, the Court said, was that it did not entail “an individual and personal injury of the kind required for Article III standing.”66 “A citizen’s interest in the overall composition of the legislature,” the Court said, “is embodied in his right to vote for his representative.”67

The Court went on to address the specific evidence before it. Among other things, the Court addressed the “efficiency gap” theory and a related “partisan symmetry” metric.68 The Court found this evidence irrelevant to individual harm:

The plaintiffs and their amici curiae promise us that the efficiency gap and similar measures of partisan asymmetry will allow the federal courts—armed with just “a pencil and paper or a hand calculator”—to finally solve the problem of partisan gerrymandering that has confounded the Court for decades. We need not doubt the plaintiffs’ math. The difficulty for standing purposes is that these calculations are an average measure. They do not address the effect that a gerrymander has on the votes of particular citizens. Partisan-asymmetry metrics such as the efficiency gap measure something else entirely: the effect that a gerrymander has on the fortunes of political parties.69

From this, the Court concluded:

That shortcoming confirms the fundamental problem with the plaintiffs’ case as presented on this record. It is a case about group political interests, not individual legal rights. But this Court is not responsible for vindicating generalized partisan preferences. The Court’s constitutionally prescribed role is to vindicate the individual rights of the people appearing before it.70

The Court, however, did not dismiss the case; it instead remanded to allow the plaintiffs another opportunity to prove standing.71

B. Back to Basics

Gill has been called a “punt,”72 and it would be that had it held simply that a plaintiff must prove residency in a district alleged to be cracked or packed in order to show standing.

But it does more. Gill refocuses the inquiry back from manageability to individual injury and therefore from party electoral success to an individual claim of right—that is, the “right to vote for his representative.” This cripples “the plaintiffs’ case as presented on this record” and the entire theory behind partisan gerrymandering. This becomes clear once we examine two basic questions about the alleged right to translate votes into seats: Who has the right? And what is the right?

1. Who Has the Right?

Gill requires partisan-gerrymandering plaintiffs to show “the effect that a gerrymander has on the votes of particular citizens.” That effect, to establish standing, cannot merely be a harm “to their interest ‘in their collective representation in the legislature,’ and in influencing the legislature’s overall ‘composition and policymaking.’”73

But there is no more to a partisan-gerrymandering claim than that alleged harm. Cracking and packing has no independent significance apart from its impact on statewide vote totals. The reason the plaintiffs in Gill complained about this practice was not that it harmed any specific voter in any specific district. The harm was that the practice had the cumulative effect of giving Republicans more, and Democrats fewer, wins across the state than their share of the vote would support in a proportional system. The district court summarized the plaintiffs’ theory in those exact terms, observing that their case depended on measuring “the proportion of ‘excess’ seats that a party secured in an election beyond what the party would be expected to obtain with a given share of the vote.”74

The missing element is the individual’s claim of right. And the Gill plaintiffs’ theory cannot simply be reworked semantically in terms of individual rights because political influence requires

---

63 Whitford, 218 F. Supp. 3d at 903.
64 Gill, 138 S. Ct. at 1921.
65 Id. at 1931.
66 Id. at 1932–33.
67 Id. at 1933 (citations omitted).
68 Id. at 1933–34.
70 Whitford, 218 F. Supp. 3d at 903–04.
concerted effort. Free-association rights, to be sure, can be described as the rights of both individuals and political parties.75 But the right asserted in a partisan-gerrymandering claim is not the right to associate or speak; it is the right to elect preferred candidates. Because an individual cannot win an election alone in a democratic system, that is a right that only can be exercised in groups. It can only be identified and enforced at an aggregate level.

Two principles in American law tie an interest in electoral influence to an individual claim of right, but both are quite different from the asserted right to translate votes into seats. The first is the one-person, one-vote rule, which requires voting districts to be substantially equal in total population to ensure that one voter’s vote does not have greater weight than another’s.76 Although it is administered at a collective level, this is an individual right because it equalizes the ratio of persons to representatives and thereby protects what Gill described as an individual’s “right to vote for his representative.”77 But the right to vote is not the same as a right to have a voter’s preferred candidate win. The right to translate votes into seats is different from the one-person, one-vote rule because it posits a right to control over who wins. Controlling outcomes can only occur by concerted action. The claim to this right therefore assumes that voters participate in the process as members of groups. The one-person, one-vote rule, by contrast, does not carry this assumption, and therefore it protects individual, not group, rights.

The second is the anti-vote-dilution principle under the Fifteenth Amendment and the Voting Rights Act (VRA). The Supreme Court has held that, if intentional, cracking and packing on the basis of race violates the Fifteenth Amendment,78 and that, even if unintentional, it violates Section 2 of the VRA.79 But both of these holdings are also founded squarely in individual rights. Both the Fifteenth Amendment and Section 2 are worded as providing individual rights to “citizens,” not groups, against the denial or abridgment of the right to vote “on account of race or color.”80 And while the VRA includes the right to an equal opportunity of minority persons “to participate in the political process and to elect representatives of their choice,” this too applies only to “members” of the racial or language-minority group, and it too is grounded in individual rights.81 What bridges the gap between the individual right to vote and the group right to influence is the immutable characteristic of race (or language-minority status). To succeed, VRA plaintiffs must show an alignment between candidate preference and racial identity by demonstrating that members of the racial group tend to vote for the same candidates. VRA plaintiffs cannot simply assume either that a voter of a particular race is likely to vote for a particular candidate or party or that a voter for a candidate or party is a member of a particular race; this correlation must be proven. VRA plaintiffs therefore must present evidence comparing these two variables to prove that a voting scheme that cracks or packs the racial group’s residents translates into a burden on their individual votes. On the other hand, if there is no proven correlation between race and candidate preference, cracking and packing has no particular meaning for individual voters because there is no way to assess the impact of their individual votes on aggregate vote totals. Accordingly, the Act ensures that “an individual’s vote will not be diluted” on the basis of an immutable and suspect characteristic.82

But a partisan-gerrymandering claim necessarily makes the kinds of assumptions VRA plaintiffs are prohibited from making, including that a group of voters identified solely by their preference for candidates—and no other shared interest or characteristic—experience an individualized harm from cracking and packing. The theory takes all voters for a specific candidate, identifies them as a group, and posits that diminished statewide vote totals harm each voter individually. This requires that every other element of a vote-dilution claim—e.g., that the group is identifiable, cohesive, and at a disadvantage as to other identifiable, cohesive groups—be assumed, as either a matter of law or a fact of political life. But few assumptions could be further removed from reality. Voters in a United States election vote for candidates, not parties. Party affiliation is only one factor among many that influence their choice. Voters routinely vote Democratic in one election and Republican in another, and many vote for Democratic and Republican candidates on the same ballot in the same election. For example, on November 7, 2000, the Pennsylvania statewide vote went to Al Gore for president and Rick Santorum for senate—two of the most polarizing political figures in each major party. Translating the harm to a political party’s vote totals to its individual voters in this context is unsupportable.

Voter preference is not like race and cannot tie the interests of a group to the interests of individuals who vote for their preferred candidates. Thus, the only way that partisan gerrymandering hurts the individual is insofar as the individual pins his or her hopes on the fortunes of the party. And, as Gill indicates, that does not establish a constitutional injury.

2. Who Does Not Have the Right?

Not only can the right to translate votes into seats be exercised only by groups, it can only be exercised by select groups: the major political parties. This is because the right to political representation is not like the rights to speech and association. Whereas allegedly harmful speech that is nonetheless constitutionally protected can always be countered by more speech, a would-be constitutional right to power for some groups can only be afforded by taking it away from others. Speech is not

75 See, e.g., Eu v. San Francisco Cty. Democratic Cent. Comm., 489 U.S. 214, 224 (1989) (“Freedom of association means not only that an individual voter has the right to associate with the political party of her choice, but also that a political party has a right to identify the people who constitute the association.”) (citations and edits omitted).

76 See Reynolds, 377 U.S. 533.

77 Gill, 138 S. Ct. at 1931.

78 See, e.g., Wesley v. Collins, 791 F.2d 1255, 1259 (6th Cir. 1986).


80 U.S. Const. amend. XV.

81 52 U.S.C. § 10301. The controversies surrounding the VRA are beyond the scope of this article.

zero-sum; party politics is. And that is true as to the competition for power both between the two major political parties and among the interest groups that combine to form them—and that could, under different circumstances, combine in alternative ways. The Republican and Democratic Parties, after all, are not facts of nature. They have formed under a specific set of circumstances, and partisan groupings would undoubtedly be different under different electoral systems.

Our political system—with its geographically-based, winner-take-all, single-member districts—is designed to favor large, big-tent parties over small, ideologically uniform parties. Notwithstanding this design, no one could or would contend that the disadvantaged interest groups have the right partisan-gerrymandering plaintiffs claim. It is implausible that, for example, pro-life conservatives could successfully petition the courts for an equal right with other groups to translate their votes into representation. No current United States electoral system provides this equal opportunity, which is why pro-life conservatives depend on the Republican Party for political influence. Aligning with the Republican Party is not their first choice; because our system renders pro-life conservatives incapable of exercising influence alone, they compromise and associate with the Republican Party to have some influence rather than none. The same can be said of environmentalists, labor-union members, libertarians, national-security hawks, and so on. As groups, they cannot translate their membership into enough votes to win elections, and their inability to do so is directly traceable to the system of representation in geographically based, winner-take-all districts. Under a different system, like many around the world, any number of different groups might vie for political power.

The asserted right to translate votes into seats, if it exists, must empower all groups to win lawsuits challenging single-member districts and winner-take-all races, since the Republican and Democratic Parties are not special constitutional creatures. And those features of our system that empower the major parties to form and exert influence infringe the supposed right of other groups to translate votes into seats at least as much as gerrymandering does. Moreover, the right of one group to translate its votes into power inevitably would run up against the same right in the hands of another group, which, due to different geographic dispersal or other characteristics of its membership, would thrive better in a different system. The right to effective influence, then, would set up a collision of the rights of virtually all American citizens against each other, given that each person will have a different view of what climate would best suit his or her chances at influence.

Identifying the injury in partisan gerrymandering is inseparable from this problem. Yes, a Republican-friendly plan diminishes Democratic Party members’ opportunity to translate their votes into representation, and vice versa. But that supposed right was already impaired for individual party members because they were compelled by practical reality to associate with each other in one of the two major parties, enormous nationwide organizations that only partially represent their views. Accordingly, this theory preferences parties over their supporters—group rights over individual rights.

By the same token, the Democratic Party could easily break a Republican-friendly gerrymander (and vice versa) by making compromises with the constituencies whose interests the gerrymander maximizes. If the districting scheme maximizes the power of suburban voters, the party can appeal to their interests; if it empowers rural voters, it can appeal to theirs. These are the same kinds of compromises that all other political groups make. And because the equal-population rule tethers representation to individual votes, a party’s chances at statewide success can never be too far divorced from its share of votes. For example, the Democratic Party in Gill complained that it would need 54% of the statewide votes to win a simple majority.83 So assuming the party could achieve 50% towards its claim of entitlement to power, the party ostensibly could defeat the gerrymander by compromising with a mere 4% of voters who previously cast votes for some Republican candidates. To be clear, the Democratic Party cannot be legally compelled to do this, but if it chooses not to, it can hardly complain that it does not control the government. Obtaining control in a democracy means responding to the system and playing the game it establishes, not manipulating it through lawsuits.

To be sure, the burdens on the Republican and Democratic Parties through gerrymandering are arguably different from the burdens on other groups insofar as gerrymandering intentionally identifies and imposes burdens on the major party out of power. But the difference is not particularly pronounced. The choice of a representational system in all cases involves a choice about which types of interests will be favored, which will not, and how they will be compelled to align in the competition for influence. The system of geographic representation and single-member districts itself is intentional—the purpose is to create a “pluralistic political process, where groups bargain among themselves” and representatives are not “beholden for office to discrete . . . groups.”84 This intentionally burdens the would-be rights of the many individuals who want purist, radical politics and representatives committed to their narrow interests or ideologies. In other words, the Democratic and Republican Parties are already benefitting from an electoral system that prioritizes their interests over competing interests. Unless the Democratic Party has rights that exceed the rights of other citizens—which is what the partisan-gerrymandering claim assumes—it has no more a constitutional right to districts that favor its interests over those of the Republican Party than those individuals have a constitutional right to a system in which the Democratic Party would cease to exist. All representational systems are created to intentionally favor certain sets of interests over others, and, as Gill holds, unless they burden individual rights, the courts have no say in how those systems are designed.

3. What Is the Right?

Gill also clarifies that a partisan-gerrymandering claim will be viable only to the extent that it asserts “individual legal rights,” which the Court distinguished from non-cognizable “generalized

---

83 Appellants’ Br., p. 9, Gill, 138 S. Ct. 1916.
84 Monroe v. City of Woodville, Miss., 881 F.2d 1327, 1329 (5th Cir. 1989) (quotations omitted).
partisan preferences.” This cuts to the heart of the partisan-gerrymandering theory, which asserts nothing other than a right to enforce partisan preference as a legal interest.

Under the Supreme Court’s vote-dilution precedents, the difference between vindicating individual rights and merely enforcing political desires depends on a plaintiff’s showing “what the right to vote ought to be.” This necessarily entails proof of “some baseline with which to compare” the challenged districting scheme. “[W]here there is no objective and workable standard for choosing a reasonable benchmark by which to evaluate a challenged voting practice, it follows that the voting practice cannot be challenged as dilutive . . . .” This is because cracking and packing inflicts an injury only to the extent it affords a group fewer seats than it otherwise would have won in a world without the injurious behavior. In other words, even if a legislature intentionally draws a map to ensure the controlling party wins X number of seats and the non-controlling party Y number of seats, there is no injury if, without the intentional gerrymandering, the controlling party would have won X number of seats and the non-controlling party Y number of seats. But how can a court identify how many seats a party would have won in a hypothetical fair election? Making this showing necessarily requires a plaintiff to prove what a “fair” districting map would be. And that is impossible because there is neither a legal standard nor a consensus anywhere on what that means.

For example, the efficiency gap theory proposes that the burden of cracking and packing should be measured against what parties “would be expected to obtain with a given share of the vote” in “a purely proportional representation system.” This necessarily assumes that the right to vote “ought to be” proportional representation. And that is so even if a legal framework predicated on the efficiency gap does not demand strict proportionality. Measuring redistricting plans against proportionality reads the assumption of proportional representation into the legal standard and measures deviations from perfection by assuming perfect proportionality as the standard. A court that imposes the efficiency gap imposes proportional representation whether or not it demands perfection, much in the same way the courts impose equality of weight in votes in the equal-population rule, even though they do not demand perfectly equal population.

But there are alternative baselines, and a court must choose which one to impose. A different baseline would be a map drawn

in accordance with “traditional districting principles,” such as compactness, contiguity, and political-subdivision integrity. The efficiency gap does not account for the values these principles may protect, and these principles do not account for any values that proportional representation may protect. Defining partisan gerrymandering by one of these baselines sets up a conflict between redistricting values. And that is exacerbated insofar as traditional districting principles are “numerous and malleable.” Even if an expert witness creates an algorithm to produce thousands of alternative maps by which to measure the alleged gerrymander, the expert necessarily plugs policy judgments into those maps by creating one algorithm and not another. “The wide range of possibilities makes the choice inherently standardless.”

All of this creates a very practical problem: the political parties’ respective jostling over the governing standard presents a severe risk that courts will engage in the very partisan gerrymandering they purport to prevent. That is because the major political parties are not similarly situated. Whereas Democratic Party voters are concentrated in cities, Republican Party voters are spread out in suburbs and rural areas. Thus, how the baseline is defined will determine whether the Constitution is read to advantage one party over another. Furthermore, because federal courts frequently must draw their own remedies to districting plans they identify as unconstitutional, a claim that a party has too difficult a task in winning seats under a plan will require courts to draw maps that assist them in winning seats. How does a court know that its “remedy” is not a gerrymander for the party that won the litigation? That is, again, an impossible question to answer because what amounts to a gerrymander in the eyes of the Republican Party is different from what amounts to a gerrymander in the eyes of the Democratic Party. There being no legal basis for choosing one baseline over another, there is no basis to choose the remedy over the invalidated plan. This problem is unavoidable because federal courts’ equitable powers are limited to correcting the legal violation, and otherwise they must “follow the policies and preferences of the State.” The partisan-gerrymandering theory makes the legal violation and the state’s “policies and preferences” indistinguishable and therefore affords federal judges no way to ascertain whether redistricting choices must be overridden (as unlawful) or followed (as legitimate state policy).

Accordingly, the very essence of a partisan-gerrymandering claim is that it vindicates partisan interests, not cognizable individual rights. Because Gill makes it clear that the Supreme Court will not recognize a claim to vindicate partisan interests, the case debilitates this cause of action.

---

85 Gill, 138 S. Ct. at 1935.
87 Id.
89 Whitford, 218 F. Supp. 3d at 904.
90 Reno, 528 U.S. at 334.
91 See Whitford, 218 F. Supp. 3d at 906 (drawing this flimsy distinction).
92 See Harris, 136 S. Ct. at 1307 (describing how federal courts enforce population equality while allowing “minor deviations from mathematical equality”) (quotations omitted). In condemning “highly disproportional representation,” the Whitford district court enforced a rule of proportional representation, even though it did not demand perfect proportionality. Allowing deviations from a principle nevertheless involves enforcing the principle.
94 Holder, 512 U.S. at 885.
96 Id.
C. Gill Going Forward

What Gill says about the scope of interests the Supreme Court is willing to protect informs, not only standing doctrine, but the political-question and equal-protection doctrines. While it remains unclear where in these potential frameworks the Supreme Court will end up, Gill, if taken at its word, makes the partisan-gerrymandering claim untenable for three reasons.

1. Gill as a Dead End: The Article III Answer

Partisan-gerrymandering claims should never proceed to the merits because no plaintiff can show individualized harm to satisfy Gill’s Article III standing rule. That is most obvious as to plaintiffs alleged to be packed into districts with fellow partisans. These plaintiffs concededly have influence over their own representatives, so they can only show harm by reference to statewide vote totals—i.e., that candidates of their preferred party were not successful in other districts, not their own. Gill holds that this is insufficient. Plaintiffs allegedly cracked into districts at levels insufficient to win in their own districts have a slightly better contention, given that their votes—assuming consistent election results over the decade, which is rare—will consistently be cast for losing candidates. But this would require the Court to find that not being represented by a member of the same party constitutes an injury to an individual.

A holding to that effect would be untenable and, indeed, damaging to democratic values. Partisan differences are typically too abstract to amount to individualized injury. They almost always concern only policy grievances about the conduct of government. Individuals vote for and against candidates based on big-picture policy questions like the national debt, foreign policy, abortion rights, judicial nominations, and so on, and not on individualized or even district-specific issues. Though undoubtedly important, public-policy issues do not create an individualized injury cognizable under Article III, so an individual can rarely claim personalized harm from being represented by a politician of the opposing party. Indeed, there frequently is little difference between Republican and Democratic candidates on localized issues specific to a district’s residents—or else they would not be competitive in the district. For example, Democratic candidates in districts with coal economies rarely inveigh against global warming; Republican candidates in districts with agricultural economies rarely campaign against farm subsidies. Only on rare occasions will a constituent be able to identify a difference with her representative that amounts to personalized harm.

Moreover, even if a plaintiff identifies such a difference, there are good reasons courts should not entertain that type of dispute. Litigating whether a representative is adequately representing a constituent would be unseemly, draw courts into political litigation to an unprecedented degree, and remove a fundamentally political question from the hands of voters and vest it in the courts. Democracy, after all, places judgment over a representative’s performance with the people. Thus, the inability of a Democratic or Republican constituent to elect a Democratic or Republican representative should not be deemed an individualized harm, and the plaintiff alleging cracking cannot otherwise identify such a harm.

Accordingly, there is no good line to draw distinguishing partisan-gerrymandering plaintiffs who have suffered harm from those who have not. They are all similarly situated in that they are claiming a restriction of rights based on statewide vote totals. If that does not confer standing—and Gill says it does not—the claims should be ruled non-viable at this threshold inquiry.

2. Vieth Revisited: Towards a Theory of Non-Justiciability

The Vieth plurality’s manageability approach, while persuasive in what it said, was deficient in what it did not say—or at least make clearer. In focusing primarily on what standards were and were not sufficiently determinate to be “manageable,” the plurality appeared to concede the underlying principle that partisan gerrymandering violates the Constitution. Commentators and lawyers could argue that, “[f]or the first time, all nine Justices agreed that excessive partisanship in redistricting is unconstitutional.” The approach was interpreted in at least some lower courts to mean that, were a clear standard identified, the claim would be viable. This was the impetus for developing social-science metrics like the “efficiency gap” that have sought to provide a precise measurement for the statewide effect of cracking and packing. But, as Gill indicates, this has largely been a red herring.

The principal problem with partisan-gerrymandering claims is not the absence of some rule of decision that is administrable in a court proceeding; the core problem is the absence of some rule tethered to the Constitution that provides a basis for courts to render what are inherently political decisions. Gill brings that latter problem into sharp focus.

This focus in Gill lays the groundwork for a more fulsome theory of justiciability. Identifying a claim as non-justiciable is not simply a matter of analyzing proposed standards for clarity; it also involves a comparison between the issue a court is asked to adjudicate and the constitutional text. But the question of which interests should and should not be favored in a redistricting, and to what degree, is inherently political, not legal. And it is entirely unrelated to the constitutional text, which says nothing about the subject. The problem is not merely that no standard is sufficiently clear or determinate; the problem is that the question is inherently standardless. Picking winners and losers in the necessary compromise of political life is a fundamentally political question that political actors, not courts, should decide.

3. Bandemer Revisited: Towards a Theory of Equal Protection

The Bandemer plurality’s opinion contains many important insights, including 1) that someone “who votes for a losing

---


98 Mitchell N. Berman, Managing Gerrymandering, 83 Tex. L. Rev. 781, 837 (2005); see also, e.g., Brief of Law Professors as Amici Curiae, p. 2. Gill, 138 S. Ct. 1916 (“The question posed by this case is not whether excessively partisan redistricting maps violate core constitutional principles. They do, and this Court has already said as much.”). That is, to be sure, a dubious reading of both the plurality and the Kennedy concurrence, since they treated the open justiciability question as a bar to deciding whether or not the Constitution was violated.

candidate is usually deemed to be adequately represented by
the winning candidate and to have as much opportunity to
influence that candidate as other voters in the district,”100 2) 
that “a failure of proportional representation alone does not
constitute impermissible discrimination under the Equal
Protection Clause,”101 and 3) that there is nothing constitutionally
problematic with “a safe district” where the plaintiff’s group
“loses election after election.”102 What it failed to do, however,
is take these principles to their plain conclusion that partisan
gerrymandering does not violate equal-protection principles.103

The flaw in Bandemer is that it assumed political parties
armed with election data were capable of proving a theory of
group rights. As experience under Bandemer showed, that has
not occurred in a single case despite dozens of attempts. Gill
sharpened this problem by requiring a theory of individual rights.
As explained above, a persuasive theory of that nature is unlikely
to be forthcoming because the reliance on election data alone
requires the assumption that voters for the same candidate in an
election are an identifiable group, such that a burden on the group
can be translated into a burden on the individual voters. But election
data alone cannot do this because it cannot link candidate preference
with some other classification; there must, at a minimum, be
some other variable in the analysis to link partisan preferences
with individual rights.

A further problem is that, if that other variable is not
a suspect classification like race, rational basis review would apply.104 Federal precedent has generally ignored legislative
motive in rational basis review cases.105 Hence, unlike in cases
alleging improper racial motive in redistricting, a partisan-
gerrymandering case would not allow inquiry beyond the text
of the redistricting statute. And that spells doom for the claim
because a redistricting statute merely “classifies tracts of land,
precincts, or census blocks.”106 There is an obvious rational basis
for those classifications,108 and it is difficult to see how, under
ordinary equal-protection principles, the claim could survive a
motion to dismiss.109

D. The Coming Dispute

There will undoubtedly emerge a competing interpretation
of Gill, which obtained unanimity only through compromise.
There is no need to speculate what that alternative view will be
because Justice Kagan offered it in her concurring opinion. The
failure, in her view, was simply an oversight by the plaintiffs’ legal
team: the plaintiffs neglected to mention at trial that they reside
in districts they believe are packed and cracked. A simple mention
of this fact would have, in her view, cured the problem.110 The
toory that cracking and packing “waste[es]” Democrats’ votes
was, in her view, perfectly valid.111

This reading is untenable for several reasons. One is that
it makes little sense of the record. The plaintiffs claimed that
all districts statewide were cracked and packed, and it was not
disputed that they lived in Wisconsin. Hence, they were claiming
that they lived in cracked and packed districts. So if the Court
would be satisfied simply with proof that the plaintiffs live in
a district alleged to be cracked or packed, it should have been
satisfied with what was before it. Instead, the decision is better
to read to hold that the plaintiffs’ burden on remand was not simply
to prove residency in cracked or packed districts; they also needed
to prove what about the cracking and packing injured them. It
is hard to make sense of the posture of the case otherwise, and
it is hard to see how they could make such a showing without
inventing an entirely new theory of the case.

A second problem with Justice Kagan’s reading is that it runs
crude against the controlling opinion’s express denial of federal-
court competency to vindicate “partisan preferences.” Cracking
and packing has practical significance only for party vote shares
and only on a statewide basis. The Court could not coherently, on
the one hand, identify injury from merely living in a cracked or
packed district, and on the other, hold that statewide injury based
on proportional vote totals is too amorphous to support standing.
A third problem with Justice Kagan’s reading is that it
suggests that a plaintiff, on the merits, can argue against all the

100 Bandemer, 478 U.S. at 132.
101 Id.
102 Id.
103 The difference between finding the claim non-justiciable and non-viable deserves further exploration. Justice Scalia’s Vrith plurality assumed that a partisan-gerrymandering claim either “presents a nonjusticiable question” or a standard that “identifies constitutional political districting.” Vrith, 541 U.S. at 350 (Scalia, J., plurality). But cases like Holder v. Hall and New York State Board of Elections v. Lopez Torres, 552 U.S. 196, 205–06 (2008), suggest a third possibility: the claim is justiciable but never viable because political districting does not violate equal-protection or free-speech rights.
104 See, e.g., City of Cleburne, Tex. v. Cleburne Living Crit., 473 U.S. 432, 440 (1985).\n105 See, e.g., Fox v. Standard Oil Co. of New Jersey, 294 U.S. 87, 100–01 (1935) (rejecting inquiry into motive in Fourteenth Amendment challenge to state taxing scheme); see also Brown v. City of Lake Geneva, 919 F.2d 1299, 1302 (7th Cir. 1990) (“[T]he motives of legislators are irrelevant to rational basis scrutiny. Instead, we must accept any justification the legislature offers for its action.”); Basket, Levy & Fine, Inc. v. St. Louis Thermal Energy Corp., 21 F.3d 237, 241 (8th Cir. 1994) (same).
107 Hunt, 526 U.S. at 547.
108 Even if motive were a permissible scope of inquiry, rational basis review requires that a statute be upheld if any rational basis can be found, so the presence of an impermissible basis does not doom a statute where a permissible basis is also present. See, e.g., FCC v. Beach Commc’ns, Inc., 508 U.S. 307, 315 (1993). There are always some permissible motives for redistricting, such as equalizing population, so virtually any redistricting statute could pass the test, even if motive were probative.
109 This was Justice Kennedy’s view of the challenge to Pennsylvania’s 2001 congressional plan, which created 13 safe Republican seats and only 5 Democratic seats in a majority-Democratic-voter state. Vrith, 541 U.S. at 313 Kennedy, J., concurring).
111 Id.
constitutional principles the Court articulated at the standing stage. A plaintiff cannot, the Court said, expect the judiciary to enforce partisan preferences; but if Justice Kagan is correct, a plaintiff, once standing is resolved, can expect the federal judiciary to enforce partisan preferences. This is not a situation, then, where the lead opinion states principle Y, offers an opportunity for the plaintiff on remand to satisfy Y, and then allows her to come back on appeal and argue for principle X. In Justice Kagan’s reading, the lead opinion states principle Y, offers an opportunity for the plaintiff on remand to satisfy principle Y, and then to come back on appeal and argue for principle not-Y.

To be sure, nothing prevents the Supreme Court from taking a contorted and illogical approach to its own precedent, and that may eventually be the result. But the lead opinion took the highly unusual step of disclaiming the concurrence, stating expressly:

Justice KAGAN’s concurring opinion endeavors to address “other kinds of constitutional harm,” perhaps involving different kinds of plaintiffs, and differently alleged burdens, see ibid. But the opinion of the Court rests on the understanding that we lack jurisdiction to decide this case, much less to draw speculative and advisory conclusions regarding others. The reasoning of this Court with respect to the disposition of this case is set forth in this opinion and none other. 112

The fact that concurring opinions carry no precedential weight is well known, so it was hardly necessary to say this. That the lead opinion went out of its way to do so is a significant red flag for anyone wishing to pursue the course set out by Justice Kagan’s concurrence. 113 It suggests that the majority of justices were well aware that the concurrence was not a concurrence, but a disguised dissent.

III. Conclusion

Gill v. Whitford is not the meaningless punt it is advertised to be. It articulates principles that undermine partisan-gerrymandering theory at the most fundamental level. Standing doctrine alone may be sufficient to solve this puzzle that has long vexed the federal courts. If nothing else, the underlying theory of rights and representation that Gill articulates, even if not fully developed, is inconsistent with partisan gerrymandering as a constitutional claim. The decision therefore should be read to definitively end these claims.

112 138 S. Ct. at 1931.

113 That Justice Kagan’s view is unlikely to prevail in the long run is further suggested insofar as any new Supreme Court justice in the mold of Justice Scalia is likely to take a formalistic approach to partisan-gerrymandering claims and look for a broad principle for resolution, either under justiciability doctrine or equal-protection law. With Justice Kennedy’s retirement, the changing makeup of the Court in the Trump era is unlikely to result in a Justice who favors a functionalist, totality-of-the-circumstances assessment of these claims. While it remains to be seen how a new Justice will approach the problem, it seems unlikely that a Trump nominee will approach it under Justice Kagan’s method.
Note from the Editor:
This article discusses the Supreme Court’s opinion in Minnesota Voters Alliance v. Mansky and criticizes the Court’s dicta suggesting that it would uphold well-written bans on political apparel in polling places.

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.


About the Author:
Michael R. Dimino is Professor of Law at Widener University Commonwealth Law School. He is a member of the executive committee of the Federalist Society’s Free Speech and Election Law Practice Group. Professor Dimino is co-author of Michael R. Dimino, Sr., et al., VOTING RIGHTS AND ELECTION LAW (2d ed. 2016) and Michael R. Dimino, Sr., et al., UNDERSTANDING ELECTION LAW AND VOTING RIGHTS (2016).


2 Minn. Stat. § 211B.11(1). Despite the potential vagueness of the term “at or about,” the parties in Minnesota Voters Alliance agreed “that the political apparel ban applies only within the polling place.” 138 S. Ct. at 1883 (emphasis in original).

3 The wearing of clothing is not literally “speech.” Nevertheless, it fits comfortably within the kind of “expressive conduct” that the Court has held to be protected by the First Amendment, such as displaying (or burning) a flag. See Texas v. Johnson, 491 U.S. 397, 402-06 (1989) (holding that burning the American flag was speech); Spence v. Washington, 418 U.S. 405, 409-11 (1974) (per curiam) (holding that the display of an upside-down American flag with a duct-taped peace sign was speech); Stromberg v. California, 283 U.S. 359 (1931) (holding that a law banning the display of a red flag violated the First and Fourteenth Amendments).

public issues. On the other hand, political speech that does not threaten those values should be protected. In particular, wearing clothing with a campaign slogan or insignia is fully consistent with the purposes of a polling place. Political-apparel bans should, therefore, be unconstitutional except as applied to conduct that presents a reasonable risk of voter intimidation or disorder.

I. INTERIORS OF POLLING PLACES ARE NONPUBLIC FORUMS

Political speech is the core concern of the First Amendment. Accordingly, regulations of political speech usually trigger strict scrutiny, requiring the government to justify such regulations by showing that they are narrowly tailored to a compelling government interest. It is indisputable that a regulation like Minnesota’s would be unconstitutional if it were applied to limit political speech on privately owned land or in a public forum. Nonpublic forums, however, are treated differently. In that category of government-owned property—places not set aside for speech or typically used to engage in debate—strict scrutiny does not apply to government restrictions on speech. Rather, speech restrictions in nonpublic forums are constitutional if they are “reasonable in light of the purpose served by the forum and are viewpoint neutral.”

The first step in determining the constitutionality of Minnesota’s political-apparel ban, then, was to determine whether the inside of a polling place should be considered a public forum. Never before had the Court considered the constitutionality of a law restricting speech within polling places. The closest precedent was Burson v. Freeman, which involved a Tennessee law that prohibited the display or distribution of campaign material and the solicitation of votes in the area around polling places. The Burson plurality applied strict scrutiny because the ban encompassed streets and sidewalks within 100 feet of polling places (an area the plurality considered to be “quintessential public forums”), but upheld the law. The plurality held that the deterrence of “voter intimidation and election fraud” was a compelling interest, and that the 100-foot “campaign-free zone” was a permissible means of achieving that interest.

If the Tennessee law was a permissible restriction of campaigning outside polling places, as Burson held, a fortiori a similar restriction on campaigning should be permissible inside polling places, where the government’s interests would be even more compelling. The Minnesota law challenged in Minnesota Voters Alliance, however, banned more speech than the Tennessee law did—including speech that was extremely unlikely to produce either voter intimidation or election fraud. Whereas the Tennessee law was principally concerned with limiting the distribution of campaign material and solicitation of votes—activities that could be hard for unwilling targets to avoid, and which might involve physical approaches and therefore a significant prospect of intimidation—the Minnesota law prohibited the merely “passive” conduct of wearing clothing with a political message or logo. It is very hard to imagine that the wearing of such clothing would lead to voter intimidation or election fraud, and even harder to imagine that apparel bans would be narrowly tailored ways of avoiding those problems. Accordingly, if political-apparel bans were evaluated under strict scrutiny, they would likely fail.

In Minnesota Voters Alliance, however, the Court held that while the traditional public forums of streets and sidewalks are often used for speech, the interiors of polling places are not: “A polling place . . . is, at least on Election Day, government-controlled property set aside for the sole purpose of voting.” Accordingly, the inside of a polling place was held to be a nonpublic forum, and the political-apparel ban triggered not strict scrutiny, but the much more lenient test applicable to nonpublic forums: whether the ban was reasonable in light of the purpose of the forum (voting) and free of viewpoint discrimination. And because the law was viewpoint neutral (at least on its face), the key question was reasonableness.

II. VAGUE BANS ON “POLITICAL” SPEECH ARE UNREASONABLE

The Supreme Court struck down Minnesota’s political-apparel ban on the narrow ground that the government had not adequately defined the kinds of apparel that were subject to the

5 See, e.g., Monitor Patriot Co. v. Roy, 401 U.S. 265, 272 (1971) (“The First Amendment has its fullest and most urgent application precisely to the conduct of campaigns for political office.”).


10 See id. at 196 (Blackmun, J., plurality opinion). By contrast, Justice Scalia, who concurred in the judgment, would have held that streets and sidewalks adjacent to polling places were nonpublic forums. See id. at 214-16 (Scalia, J., concurring in the judgment).

11 See id. at 206 (plurality opinion).

12 See id. at 210.

13 See Minnesota Voters Alliance, 138 S. Ct. at 1887 (“The [Burson] plurality’s conclusion that the State was warranted in designating an area for the voters as ‘their own’ as they enter the polling place suggests an interest more significant, not less, within that place.”) (emphasis omitted).

14 It also prohibited the “display of campaign posters, signs or other campaign materials,” but the ban on displays was not specifically discussed by the Burson plurality.

15 Minnesota Voters Alliance, 138 S. Ct. at 1887.

16 Id. at 1886.

17 See Minnesota Voters Alliance, 138 S. Ct. at 1886 (‘A polling place in Minnesota qualifies as a nonpublic forum.”).
prohibition. The ban was therefore too “indeterminate” and likely to lead to “erratic application.”

Minnesota’s political-apparel ban provided that a “political badge, political button, or other political insignia may not be worn at or about the polling place.” Election judges at each polling place were authorized to determine if a particular piece of clothing fell within the prohibition, and to refer the matter to other officials if the offending voter refused to remove the item. Offenders were allowed to vote, but they were subject to a fine of up to $300.

The challengers who brought the case to the Supreme Court were individuals and associations of individuals who wished to wear clothing promoting the Tea Party and buttons stating “Please I.D. Me.” When they attempted to vote, some of the challengers were asked to cover up their political apparel, and those who refused to do so had their names recorded for referral and possible prosecution. One of the challengers was twice barred from voting until he removed or covered up the political apparel, despite the fact that the Minnesota law did not permit election officials to turn away voters who persisted in wearing political apparel.

In an attempt to clarify the meaning of “political”—and to make clear that the statute did not ban all speech having to do with government, politics, or elections—the Minnesota secretary of state issued a guidance document to the state’s election officials, stating that the prohibited political apparel “includes, but [is] not limited to” any item containing “the name of a political party” or “the name of a candidate at any election”; “[a]ny item in support of or opposition to a ballot question at any election”; “[i]ssue-oriented material designed to influence or impact voting (including specifically the ‘Please I.D. Me’ buttons)”; and “[m]aterial promoting a group with recognizable political views (such as the Tea Party, MoveOn.org, and so on).” Overall, Minnesota argued that the ban did not apply to all material that might be considered “political,” but rather to material having to do with the choices faced by the voters at the election.

The Court was troubled by the vagueness of the ban, even as clarified by the secretary of state. The government wished to apply the ban on issue-oriented material to issues that had been raised in the election campaigns. But, as the Court pointed out, “[a] rule whose fair enforcement requires an election judge to maintain a mental index of the platforms and positions of every candidate and party on the ballot is not reasonable.” The issue-related ban also led to some seeming inconsistencies. For example, at oral argument, the state represented that a voter could not wear a shirt that contained the text of the Second Amendment, but could wear one that contained the text of the First Amendment.

The ban on materials associated with groups holding “recognizable political views” was hopelessly vague. As the Court noted, “[a]ny number of associations, educational institutions, businesses, and religious organizations could have an opinion on an ‘issue[,] confronting voters in a given election.’” The state’s attempt to limit the ban to groups holding views that were “well-known” to the “typical observer” raised more problems, for the application of those standards “may turn in significant part on the background knowledge and media consumption of the particular election judge.” For these reasons, the Court concluded that the statute was unconstitutionally vague.

Even the two dissenters did not disagree that the ban’s vagueness created constitutional concerns. Justice Sotomayor, joined by Justice Breyer, urged the Court to certify the case to the Minnesota Supreme Court so that the state court could construe the ban to avoid the law’s apparent vagueness and arbitrary distinctions. In the dissenters’ view, the state court’s construction “likely would obviate the hypothetical line-drawing problems” identified by the majority of the Supreme Court.

III. Would a Well Written Ban on Political Apparel Be Reasonable?

There were two ways that the Minnesota political-apparel ban could have been held to fail the test of reasonableness applicable to speech restrictions in nonpublic forums. The first way, adopted by the Court, was to hold that the law was unreasonable because it was vague. Potential speakers could not determine what messages would be determined to be “political,” and the vagueness of the law meant that the officials charged with its enforcement had too much discretion—which could have been used to discriminate against disfavored viewpoints. The alternative approach would have addressed a more fundamental objection to the ban: It simply prohibited too much speech—far more than necessary to serve the government interests at issue.

Minnesota Voters Alliance suggested in dicta that a “more lucid” statute than Minnesota’s would have been constitutional, and it pointed to statutes in California and Texas as examples of the kinds of restrictions that states may constitutionally adopt. Those laws ban political apparel in polling places, but only political apparel that references a candidate, a ballot measure, or (in Texas) a political party. While the California and Texas bans are more clearly written than the Minnesota one, they still ban political speech. Thus, according to the public-forum doctrine discussed above, they would have to be reasonable and viewpoint

18 Id. at 1891.
19 Id. at 1890.
20 Minn. Stat. § 211B.11(1).
21 See Minnesota Voters Alliance, 138 S. Ct. at 1883.
22 Id.
23 Id. at 1889.
24 Id.
25 See id. at 1891 (quoting Tr. of Oral Arg. at 40).
26 Id. at 1890.
27 Id.
28 Id. at 1893 (Sotomayor, J. dissenting).
29 See id. at 1891 (opinion of the Court) (offering the California and Texas statutes as examples or “more lucid” laws, but stating that “[w]e do not suggest that such provisions set the outer limit of what a State may proscribe, and do not pass on the constitutionality of laws that are not before us”).
neutral to be constitutional. Although the Court seemed inclined to view such statutes as constitutional, it should have been more cautious about saying so. Such bans are likely unreasonable, and therefore unconstitutional, because they restrict far more speech than necessary to serve the government's interests in conducting elections and permitting voters to cast votes in an appropriately contemplative atmosphere.

Assessing reasonableness first requires one to identify with particularity the government interests a speech restriction is supposed to serve. Burson accepted that the government has a compelling interest in avoiding voter intimidation and electoral fraud. But there is no reason to think banning political apparel would have any effect on electoral fraud. And as to voter intimidation, the voter who is intimidated by another voter's T-shirt is unusually susceptible to influence;35 the vast majority of voters would be no more intimidated by another voter's clothing than by a neighbor's yard sign. Viewed from the speaker's perspective, on the other hand, political-apparel bans impose a burden on voting. If a voter wants to display his political apparel outside the polling place, he may have to bring a change of clothes for when he goes inside. If he forgets extra clothes, he may have to go home, return to the polling place, and wait in line again, before voting. If the polls have closed in the meantime, the ban could cost the voter the opportunity to cast a ballot. The political-apparel ban thereby imposes a non-trivial burden on speech and voting rights and barely, if at all, serves the compelling interest in preventing voter intimidation. The ban is therefore an unreasonable way of achieving those interests.

In nonpublic forums, however, the government has authority to pursue interests beyond those that are compelling. Indeed, in a nonpublic forum, the government may impose speech restrictions that serve no other interest than preserving the forum for its non-speech purpose.31 Simply stated, the government may restrict speech inside a polling place if the restriction is a reasonable way of preserving the polling place as a location for voting. While such a test is deferential to the government (and appropriately so), not all bans on speech in nonpublic forums are reasonable. Sometimes the restrictions sweep too broadly, and the First Amendment protects speakers whose expressive conduct presents little risk of interfering with the purpose of the forum—in this case, enabling voters to cast votes free of pressure or conflict.

Surely states may ban polling-place speech that actually makes it difficult to cast or record votes, that threatens other voters, or that “is intended to mislead voters about voting requirements and procedures.”32 Just as surely, the government may not ban certain speech simply because it would rather designate an area as off-limits to that kind of speech. Thus, it is crucial to specify what it means to preserve polling places for voting. Minnesota Voters Alliance phrased the government interest variously as “set[ting] a polling place aside as ‘an island of calm in which voters can peacefully contemplate their choices’”;33 “reflect[ing] th[e] distinction between ‘choosing [and] campaigning’;”34 “ensur[ing] that partisan discord not follow the voter up to the voting booth, and distract from a sense of shared civic obligation at the moment it counts the most”;35 and allowing voters to “focus on the important decisions immediately at hand.”36

States may justifiably work to promote voters' “peaceful contemplat[ion]” and “focus.” Any activity that disrupts an individual's ability to cast a vote for a chosen candidate, or that prolongs the voting process by interfering with voters' ability to concentrate, interferes with the voting process itself. It is by no means clear, however, that states may pass speech restrictions with the purpose of suppressing “partisan discord” or “campaigning” unless there is some other reason that discord or campaigning is harmful, such as interfering with voters' free choices or, perhaps, “distract[ing] from a sense of shared civic obligation.”36 Discord itself, the Court has long recognized, is a natural by-product of the First Amendment, which expresses a “profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks on government and public officials.”37 Partisan discord, in other words, advances self-government unless there is something

30 Cf. Doe v. Reed, 561 U.S. 186, 228 (Scalia, J., concurring in the judgment) (“There are laws against threats and intimidation; and harsh criticism, short of unlawful action, is a price our people have traditionally been willing to pay for self-governance. Requiring people to stand up in public for their political acts fosters civic courage, without which democracy is doomed.”).
31 If the purpose of the forum were speech, the forum would not be a nonpublic forum. Rather, it would be a designated public forum, and speech restrictions would be evaluated under strict scrutiny. See Good News Club v. Milford Central School, 533 U.S. 98 (2001); Rosenberger v. Rector and Visitors of the University of Virginia, 515 U.S. 819 (1995).
32 Minnesota Voters Alliance, 138 S. Ct. at 1889 n.4.
33 Id. at 1887 (quoting respondents' brief at 43).
34 Id. at 1888.
35 Id.
36 Id. It is difficult to know what this means, and even more difficult to understand how political apparel could provide such a distraction. The Court might mean that voting should be a solemn act and that voters' clothing should reflect that attitude, but it is impossible to square that interest with a law that bans political apparel but permits people to vote in tank-tops and flip-flops. More likely, the Court means that the state should be able to promote a united front—the appearance that all voters are at the polling place “to reach considered decisions about their government and laws” in a non-partisan manner. Such an interest is farcical. Voters are there to choose one candidate over another, on (for most races, at least) partisan ballots. For the state positively to promote partisanship by printing partisan ballots and then to restrict speech based on the pretense that voters should not act in a partisan manner at the polling place is absurd.
about the way the speech is expressed that interferes with the voting process. Some speech within the polling place surely can produce such interference, but not all speech will, and it is the interference with voting that might be caused by the speech, rather than the discord itself, which is the harm that states should try to prevent. States have no interest in avoiding public disagreement per se; such an interest seems to be nothing less than aversion to the clash of views that is an inherent (not to say beneficial) part of free debate protected by the First Amendment. 38 As the Court eloquently said in *Texas v. Johnson* in reversing a conviction for flag-burning, “If there is a bedrock principle underlying the First Amendment, it is that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.” 39 Thus, states should not be able to create campaign-free zones if the only reason for doing so is that the government would prefer that campaign speech not occur in that place, or that some people encountering the speech will disagree with it; rather, the government should have to point to some reason that speech in that location would cause harm. It is particularly doubtful that a state could have a legitimate interest in suppressing “partisan discord,” given that (as the Court recently reaffirmed in *Reed v. Town of Gilbert, Arizona*) 40 discrimination between different content-based categories of speech triggers strict scrutiny and political speech is the kind of speech most central to the purpose of the First Amendment. Thus, speech restrictions in nonpublic forums may be reasonable if they promote voters’ free choices and calm reflection, but not if they exclude political speech from polling places despite the lack of any threat to those interests.

Measured by this standard, most political-apparel bans are unconstitutional. States should be able to ban active political speech in polling places—for example, approaching or addressing other voters to persuade them to support a candidate or ballot measure. *Burson* recognized states’ interest in protecting voters from a barrage of campaigning immediately outside of the polling place, and that interest is even stronger inside the building. 41 Within the polling place, states should be able to prohibit all loud communication, and perhaps all oral communication unconnected with voting, so as to preserve the peace and quiet that facilitate reflection by voters. Voters in the polling place are a captive audience, and states should be able to protect them from unwanted noise. 42 States might even be able to discriminate by content and prohibit oral communication about campaigns for candidates and ballot measures, on the ground that such communication could present the greatest interference with other voters’ ability to decide on those very candidates and issues. Passive displays of clothing, however, would not interfere with other voters’ ability to reflect on their choices and make their selections. A voter concentrating on his ballot may be disrupted by a voice; he will not be disrupted by another voter’s T-shirt. One can turn away from a visual display and limit the distraction. 43 Sounds, however, are far more invasive. So long as one is within earshot of a distracting sound, that sound can force its way into one’s consciousness and interfere with the ability to concentrate on other tasks. Once one looks away from a visual display, however, it is distracting only as a memory. The memory of seeing another voter’s T-shirt is thus little different from the memory of seeing a piece of political apparel—or a yard sign, or a billboard—outside the polling place, whereas distracting sounds present a continuing bombardment of our consciousness whether we direct our attention to them or not.

In the analogous context of government workplaces, the Court has protected expression so long as it does not obstruct government functions. 44 And the Court has not simply rubber-stamped the government’s claims that speech would lead to obstruction of its functions. In *Pickering v. Board of Education* and *Rankin v. McPherson*, for example, the Court held that the government’s interests in the effective operation of schools and law-enforcement agencies did not require it to restrict employees’ speech. *Pickering* upheld a teacher’s right to publish a letter critical of his school board, and *McPherson* protected the right of an employee in a constable’s office to make a remark supporting the assassination of President Reagan. In both cases, the Court demanded that the government show that the speech would harm the functioning of the government office, and the Court analyzed the facts of each case before concluding that no such harm was likely. 45 By analogy, the Court should not blindly defer to states’ claims that all political apparel presents a risk of interference with other voters’ ability to vote.

Schools are another nonpublic forum in which speech restrictions are often challenged, and they may be the most analogous context to polling places. In both contexts, the government must carry out a function (education/running an election) other than providing a forum for speech. In both schools and the polls, an excessive amount of speech (e.g., raucous chanting) could interfere with that government function. Neither schools nor polling places exist as forums for speech, yet both education and voting depend on the exchange of ideas that is protected by the First Amendment. Further, in both contexts, bans on political apparel would substantially limit speakers’ ability to reach their intended audiences (classmates/other voters). If schoolchildren were to be allowed to display political messages only outside of the school, and voters were allowed to display political messages only outside of polling places, each group of speakers would have less of an opportunity to communicate their

---

38 See, e.g., *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting) (classically expressing the marketplace-of-ideas theory behind the First Amendment). See also, e.g., *Snyder v. Phelps*, 562 U.S. 443, 461 (2011) (protecting the speech of funeral protestors, and noting that “even hurtful speech on public issues” must be protected “to ensure that we do not stifle public debate”).

39 491 U.S. at 414.


41 See Minnesota Voters Alliance, 138 S. Ct. at 1887.


43 See *Cohen*, 403 U.S. at 21 (noting that persons offended by the words on Cohen’s jacket “could effectively avoid further bombardment of their sensibilities simply by avert[ing] their eyes”).


messages to their intended audiences. Such a limitation should be permissible only where the communication of the messages has a realistic chance of producing harm.

Minnesota Voters Alliance seemed content to resolve this conflict by saying that the time for speech was before one arrived at the polls: “Casting a vote . . . is a time for choosing, not campaigning.” 46 Yet the Court adopted a markedly less restrictive approach with respect to schools. In Tinker v. Des Moines Independent Community School District, 47 the Court held that the school could not prohibit students from wearing arm bands that protested the Vietnam War, and it rejected Justice Hugo Black’s argument that school officials should be able to exclude political speech from classrooms so that students could “keep their minds on their own schoolwork.” 48 Although the Court recognized that speech could be disruptive to schools’ ability to educate students, the Court was adamant that the mere possibility of such disruption was insufficient to justify a limitation on speech, even within schools: “[U]ndifferentiated fear or apprehension of disturbance is not enough to overcome the right to freedom of expression.” 49 The Tinker Court demanded record facts to support the school’s contention that the arm bands would disrupt the functioning of the school, and it found no such support. The protesting students “neither interrupted school activities nor sought to intrude in the school affairs or the lives of others. They caused discussion outside of the classrooms, but no interference with work and no disorder.” 50 The parallel with polling-place speech is apparent: Wearing “political” apparel—be it ideological, partisan, or candidate-specific—need not cause any interference with voting or disorder.

Minnesota Voters Alliance recognized the potential parallel with Tinker and further noted that in another case the Court had characterized the wearing of political apparel in an airport as “nondisruptive.” 51 The Court attempted to dismiss the relevance of these precedents by asserting that neither case involved the purportedly “unique context of a polling place on Election Day.” 52 The Court did not, however, explain what was so different about polling places that a speech restriction could be justified there, but would be unconstitutional in schools. 53 The two contexts seem remarkably similar; in both instances the government’s important function could suffer from noisy or distracting political speech, but not from mere displays of political ideology or affiliation on apparel or accessories. If schoolchildren can be trusted to ignore their classmates’ political apparel and concentrate on their lessons, we should be able to trust adults to ignore others’ political apparel and concentrate on voting for the few minutes that they are in the voting booth. 54 In fact, speech by adult voters should be even more protected than speech by schoolchildren, as school authorities are permitted much greater control of students’ lives than election officials are able to exercise over voters. 55 Further, children are more impressionable and distractible than are adults, so speech in schools is more likely to interfere with schools’ educational function than political apparel is to interfere with the ability to vote. Just as “[n]either students [n]or teachers shed their constitutional rights to freedom of speech or expression at the schoolhouse gate,” 56 voters should not be deemed to shed their constitutional rights to freedom of speech or expression at the entrance to the polling place.

IV. Conclusion

Minnesota’s political-apparel ban not only restricted political speech, but it was unclear about what speech would be treated as “political.” The state’s interpretations, meant to clarify the scope of the ban, created arbitrary distinctions and drew lines that were nearly impossible to administer. Accordingly, the Supreme Court easily concluded that the ban was unconstitutional without a clearer definition of the kind of “political” apparel that was barred from the polling place.

More significant for the future are the Court’s dicta suggesting that if states clearly identify the apparel that is prohibited, they may impose polling-place bans on clothing containing political messages or logos. These dicta go too far in approving restrictions of political speech. States need to ensure that polling places provide an environment that allows voters to think and concentrate on the choices they are making, and to make those choices without undue influence or intimidation. Wearing a “Make America Great Again” hat or a “Yes We Can” shirt, however, does nothing to undermine the purpose of a polling place because it does not intimidate voters or interfere with voters’ ability to contemplate the questions on the ballot.

The Court was correct to hold that polling places are nonpublic forums. Accordingly, the government may restrict speech within polling places so long as those speech restrictions are reasonable and viewpoint neutral. “Reasonable,” however, does not mean that the government has carte blanche to restrict

---

46 138 S. Ct. at 1887.
48 Id. at 518 (Black, J., dissenting).
49 Id. at 508 (majority opinion).
50 Id. at 514.
51 138 S. Ct. at 1887 (quoting Board of Airport Commis’rs of Los Angeles, 482 U.S. at 576).
52 Id. at 1887.
53 The Court instead drew a parallel between polling places and courtrooms or legislatures, arguing that “[c]asting a vote is a weighty civic act, akin to a jury’s return of a verdict, or a representative’s vote on a piece of legislation.” Id. The comparison is peculiar if the Court wants to justify some political-apparel bans in polling places, because surely legislators cannot be prohibited from wearing political apparel or making political speeches in the legislature.
54 Minnesota Voters Alliance’s suggestion that a content-based speech restriction at a polling place could enable voters to “focus on the important decisions immediately at hand,” id. at 1888, is in some tension with Tinker’s conclusion that the ban on arm bands was not necessary to enable students to focus on their studies. Even assuming that it is more important for voters than for students to focus, speech restrictions should be appropriate only when there is a significant risk that the speech will actually cause voters to lose focus.
55 See, e.g., Morse v. Frederick, 551 U.S. 393 (2007) (permitting school officials to punish a student for expressing a pro-drug message at a school-sponsored event).
56 Tinker, 393 U.S. at 506.
expression whenever it would prefer not to see it. Rather, there must be some *reason* to think that the expression would interfere with the ability of people to cast their votes. The Court should recognize that one voter’s clothing does not interfere with another voter’s rights. Our democracy accepts—and in some ways depends on—our differences of opinion about politics. Far from being a threat to democracy, our political differences make democracy meaningful. So long as those differences are expressed in a way that does not intimidate others, we should be proud to live in a country that celebrates our ability to vote and to voice our opinions.
Compelled Speech in Masterpiece Cakeshop: What the Supreme Court’s June 2018 Decisions Tell Us About the Unresolved Questions

By James A. Campbell

Note from the Editor:

This article discusses the unresolved compelled-speech questions in Masterpiece Cakeshop v. Colorado Civil Rights Commission. It argues that the Court hinted at how it will ultimately resolve those questions in the various Masterpiece Cakeshop opinions and in its opinions in Janus v. AFSCME and NIFLA v. Becerra.

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.


Many legal commentators expected the Supreme Court of the United States to make a big splash on freedom of speech in the summer of 2018. Most eyes were focused on Masterpiece Cakeshop v. Colorado Civil Rights Commission. So when early June 2018 brought a decision in Masterpiece that focused on religious exercise, many assumed that the Court simply withheld guidance on the hotly debated compelled-speech questions raised in that case.

But tucked amidst the Court’s free-exercise analysis, Masterpiece provides insight into how the speech question should be resolved. And later in June, the Court issued two other decisions addressing compelled speech: National Institute of Family and Life Advocates (NIFLA) v. Becerra and Janus v. American Federation of State, County, and Municipal Employees, Council 31 (AFSCME). Together, these three decisions—Masterpiece, NIFLA, and Janus—support the right of creative professionals to decline to create speech or artistic expression that violates their conscience.

Part I of this article discusses the background of the Masterpiece case. Part II explores the various opinions issued in Masterpiece. Part III provides a brief overview of key portions of the majority opinions in NIFLA and Janus. Part IV identifies two speech-related issues—one statutory and one constitutional—that courts must resolve after Masterpiece, and it discusses two ongoing cases that illustrate the contours of those issues. Finally, Parts V and VI analyze how relevant portions of Masterpiece, NIFLA, and Janus point toward a resolution of the compelled-speech questions that will enable creative professionals to make a living without violating their consciences.

I. Masterpiece Cakeshop

Jack Phillips, the owner of Masterpiece Cakeshop, is a cake artist. He uses his skills as a pastry chef, sculptor, and painter to create works of art in the form of elaborate, custom-designed cakes. The crown jewels of Phillips’s artistry are his custom-designed wedding cakes. Phillips is also a man of deep religious faith whose beliefs guide his work. Those convictions inspire him to serve people from all walks of life, but to decline to create cakes that express messages or celebrate events in violation of the tenets of his faith. His decisions on whether to design a custom cake have never focused on who the customer is, but on what the custom cake will express or celebrate.

In the summer of 2012, a same-sex couple entered Masterpiece Cakeshop to discuss a wedding cake celebrating their marriage. Phillips told the gentlemen that he could not create such a cake, but that he would sell them anything else in his shop or design a cake for them for a different occasion. Phillips does

About the Author:

James A. Campbell serves as senior counsel with Alliance Defending Freedom, where he is director of the Center for Cultural Engagement and Scholarship. He represents Jack Phillips in Masterpiece Cakeshop v. Colorado Civil Rights Commission.
not create wedding cakes celebrating same-sex marriage because doing so would require him to express through his art celebration for a view of marriage that conflicts with his religious beliefs.8

The two men filed complaints with the state of Colorado, alleging that Phillips discriminated against them because of their sexual orientation.9 Phillips argued that he did not turn the men away because of their sexual orientation, but because of the message that he would have communicated through a wedding cake celebrating a same-sex marriage. Phillips also contended that the government could not punish him under these circumstances because doing so would violate his freedoms of religion and speech guaranteed under the First Amendment.10

After an investigation and hearings, the Colorado Civil Rights Commission (Commission) found that Phillips had engaged in unlawful discrimination.11 It ordered him to do three things.12 First, he had to either start designing cakes celebrating same-sex weddings or stop creating wedding cakes altogether. Second, he had to implement staff training on compliance with nondiscrimination law, which would require him to tell his staff, most of whom are his family members, that he was wrong to decline requests to design custom wedding cakes celebrating same-sex marriages. Third, he had to submit periodic reports disclosing the cake requests he declined and explaining the reasons.

Soon after the Commission issued its order, the commissioners discussed Phillips’s case again at one of their public hearings. During that meeting, the commissioners expressed outright hostility toward Phillips’s claim that he should be free to create his custom cake art consistently with his faith. One commissioner, with no objection from the others, said:

I would also like to reiterate what we said in the hearing or the last meeting. Freedom of religion and religion has been used to justify all kinds of discrimination throughout history, whether it be slavery, whether it be the Holocaust, whether it be—I mean, we—we can list hundreds of situations where freedom of religion has been used to justify discrimination. And to me it is one of the most despicable pieces of rhetoric that people can use to—to use their religion to hurt others.13

Phillips appealed the Commission’s order to the Colorado Court of Appeals, which affirmed it.14 In its opinion, the court acknowledged that Phillips declined the cake request because of his beliefs about marriage rather than his “opposition to [the customers’] sexual orientation.”15 Nonetheless, the court reasoned that the state public-accommodation law requires no “showing of ‘animus’” against individuals16 and held that Phillips violated the statute simply by declining “to create a wedding cake for [the customers’] same-sex wedding celebration.”17 As part of that analysis, the court of appeals distinguished three other cases—decided around the same time—in which the Commission found no discrimination when three cake artists refused a religious man’s requests for cakes criticizing same-sex marriage on religious grounds.18 Those other cake artists, the court explained, “did not refuse the patron’s request because of his [religion], but rather because of the offensive nature of the requested message.”19 The court of appeals also rejected all of Phillips’s First Amendment claims. Concerning his free-speech claim, the court held that Phillips “does not convey a message supporting same-sex marriages merely by abiding by the law” because “a reasonable observer would understand that [his] compliance with the law is not a reflection of [his] own beliefs.”20

After the Colorado Supreme Court declined to hear his case, Phillips raised two issues to the Supreme Court of the United States. The first was his free-speech claim that the government violated his expressive freedom by requiring him to create custom wedding cakes celebrating same-sex marriage. Phillips argued that the First Amendment’s protection against compelled speech, which should shield him in this case, applies when two factors are satisfied: (1) when a customer asks a creative professional to create a custom work that qualifies as constitutionally protected expression; and (2) when the professional declines the request because of the message that the custom work would communicate rather than simply because of who the customer is.21

The state, in contrast, argued that the compelled-speech doctrine offers no protection when governments apply public-accommodation laws to people who earn a living by creating and selling expression. Under this view, whether Phillips’s custom wedding cakes qualify as speech is irrelevant because, even if they do, the First Amendment affords him no relief. The state argued that this rule will not result in widespread compulsions of speech because governments apply public-accommodation laws to speech only when expressive professionals decline to create for some the same words, designs, or messages that they have created for others.22

Phillips based his second claim on the Free Exercise Clause, arguing that the Commission violated his religious freedom by manifesting hostility toward his faith and by treating his religious...
decision not to celebrate same-sex marriage worse than the decisions of other cake artists not to criticize same-sex marriage.

II. The *Masterpiece* Opinions

The Supreme Court ruled 7-2 in Phillips’s favor. It did so exclusively on free-exercise grounds. Because the government’s hostility toward Phillips’s faith was so apparent, the Court did not need to reach the free-speech question.

A. The Majority Opinion

The majority opinion held that the Commission displayed “clear and impermissible hostility toward [Phillips’s] sincere religious beliefs” about marriage and that it therefore violated his right to free exercise of religion. It pointed to two ways the Commission displayed hostility. For one, the Court found an “indication of hostility [in] the difference in treatment between Phillips’[s] case and the cases of other bakers who object[ed] . . . on the basis of conscience” to requests for “cakes with images that conveyed disapproval of same-sex marriage.” While the Commission punished Phillips for following his conscience, it gave the other cake artists a pass. The Commission “found no violation . . . in the other cases in part because each bakery was willing to sell other products . . . to the prospective customers.”

But it “dismissed Phillips’[s] willingness to sell [other items] to gay and lesbian customers as irrelevant.”

In addition, the Court discerned hostility “at the Commission’s formal, public hearings.” The majority highlighted a number of comments that “show[ed] lack of due consideration about marriage and that it therefore violated A. The Majority Opinion

not need to reach the free-speech question.

The combination of unequal treatment and hostile remarks left no doubt that the government failed to consider Phillips’s religious claims “with the neutrality that the Free Exercise Clause requires.” Although the Court did not reach the free-speech question, it provided lower courts with some guidance for resolving cases that raise a “confluence of speech and free exercise principles.” It admonished courts to balance respect for the fundamental First Amendment freedoms of religious adherents who believe that marriage is the union of a man and a woman with respect for the dignity of LGBT individuals. Notably, the Court did not reject or foreclose any free-speech argument that Phillips raised in the case. In fact, as explained in Section VI, the free-speech roadmap that the *Masterpiece* majority laid out is consistent with the position that Phillips advocated.

B. Justice Kagan’s Concurrence

Justice Kagan, who joined the majority, also authored a concurrence. She explained that the state’s actions were particularly “disquieting” because an “obvious” basis existed for “distinguishing” Phillips’s case from the cases involving the other three cake artists. While Phillips declined to create “a wedding cake that [he] would have made for an opposite-sex couple”—one “suitable for use at same-sex and opposite-sex weddings alike”—the other cake artists, Justice Kagan wrote, declined “to make a cake . . . that they would not have made for any customer.” In refusing that request, the bakers did not single out that customer because of his religion, but instead treated him in the same way they would have treated anyone else—just as the public-accommodation law requires. In short, she said, a “vendor can choose the products he sells, but not the customers he serves.”

Justice Kagan thus adopted a variation of the state’s argument that business owners violate a public-accommodation law only when they decline to create for one person an item containing the same words, symbols, and messages that they created for another.

C. Justice Gorsuch’s Concurrence

Justice Gorsuch also wrote a concurring opinion, which Justice Alito joined. That opinion explained thatPhillips’s case and those of the bakers who refused to make cakes opposing same-sex marriage “share all legally salient features.”

[T]here’s no indication the bakers actually intended to refuse service because of a customer’s protected characteristic.

---

23 *Masterpiece*, 138 S. Ct. at 1729.
24 Id. at 1730.
25 Id.
26 Id.
27 Id. at 1729.
28 Id.
29 Id.
30 Id. See supra note 13.
31 *Masterpiece*, 138 S. Ct. at 1731.
32 Id. at 1723.
33 See id. (“The case presents difficult questions as to the proper reconciliation of at least two principles. The first is the authority of a State and its governmental entities to protect the rights and dignity of gay persons who are, or wish to be, married but who face discrimination when they seek goods or services. The second is the right of all persons to exercise fundamental freedoms under the First Amendment, as applied to the States through the Fourteenth Amendment.”); id. at 1732 (“The outcome of cases like this in other circumstances must await further elaboration in the courts, all in the context of recognizing that these disputes must be resolved with tolerance, without undue disrespect to sincere religious beliefs, and without subjecting gay persons to indignities when they seek goods and services in an open market.”).
34 Id. at 1733 (Kagan, J., concurring).
35 Id.
36 Id. at 1733 n.*.
37 Id. at 1733; see also id. at 1733 n.* (explaining that those three cake artists “would not sell the requested cakes to anyone”).
38 Id. at 1733.
39 Id. at 1733 n.*.
40 Id. at 1735 (Gorsuch, J., concurring).
We know this because all of the bakers explained without contradiction that they would not sell the requested cakes to anyone, while they would sell other cakes to members of the protected class (as well as to anyone else). . . . [I]t was the kind of cake, not the kind of customer, that mattered to the bakers.41

Justice Gorsuch disagreed with Justice Kagan’s view that Phillips’s case is distinguishable from the cases brought against the three other cake artists. In particular, he objected to (1) Justice Kagan's characterization of Phillips’s case as involving “wedding cakes”—and not a wedding cake celebrating a same-sex wedding—and (2) her supposition that all wedding cakes are “indistinguishable.”42 He said that by focusing on wedding cakes instead of wedding cakes celebrating same-sex marriage, both the state and Justice Kagan played with “the level of generality”—“adjusting the dials just right.”43 He considered this an “improper” kind of “results-driven reasoning” that “risks denying constitutional protection to religious beliefs that draw distinctions more specific than the government’s preferred level of description.”44

D. Justice Thomas’s Concurrence

Justice Thomas wrote another concurrence, which Justice Gorsuch joined. His is the only opinion that squarely addressed the free-speech issue. While public-accommodation laws are constitutional in most of their applications, Justice Thomas explained, “the First Amendment applies with full force” when those laws declare “speech itself to be the public accommodation.”46 Here, Phillips’s wedding cakes “do, in fact, communicate” that “a wedding has occurred, a marriage has begun, and the couple should be celebrated.”47 “If an average person walked into a room and saw a white, multi-tiered cake, he would immediately know that he had stumbled upon a wedding.”48 “Forcing Phillips to make custom wedding cakes for same-sex marriages requires him to, at the very least, acknowledge that same-sex weddings are ‘weddings’ and suggest that they should be celebrated—the precise message he believes his faith forbids.”49

Justice Thomas also explained why the state’s reliance on the “dignity” of potential customers does not justify violating Phillips’s free-speech rights.50 Such “justifications are completely foreign to [the Court’s] free-speech jurisprudence.”51 “States cannot punish protected speech because some group finds it offensive, hurtful, stigmatic, unreasonable, or undignified.”52 The government thus cannot force Phillips’s artistic expression in order to protect others’ dignity.

Justice Thomas concluded by emphasizing that in future cases, the freedom of speech could be essential to preventing the government from “vilify[ing] Americans who are unwilling to assent to the new orthodoxy” on marriage.53 That freedom must “maintain its vitality.”54

E. Justice Ginsburg’s Dissent

Justice Ginsburg, together with Justice Sotomayor, dissented. They disagreed with the majority’s conclusion that “Phillips’s religious objection was not considered with the neutrality that the Free Exercise Clause requires.”55 Beginning with the unequal treatment between the various cake artists, Justice Ginsburg regarded the cases as “hardly comparable.”56 The bakers would have refused to make a cake with [the religious customer’s] requested message for any customer, regardless of his or her religion. And the bakers . . . would have sold [that customer] any baked goods they would have sold anyone else.57 That customer, in other words, “was treated as any other customer would have been treated—no better, no worse.”58 In contrast, Justice Ginsburg said, Phillips refused to design “a cake of the kind he regularly sold to others.”59 The dissenters placed great weight on the distinction between declining a custom item “with a particular design and whose form was never even discussed.”60

Furthermore, Justice Ginsburg saw “no reason why the comments of one or two Commissioners should be taken to overcome” what she viewed as Phillips’s unlawful actions.61 She emphasized that the Colorado proceedings involved “several layers” of decision-making: (1) the Colorado Civil Rights Division, (2) an administrative law judge, (3) the Commission, and (4) the Colorado Court of Appeals.62 Because she discerned no prejudice outside of the commissioners’ hostile comments, she did not think that bias during that part of the proceedings could render Phillips’s punishment unconstitutional.63

51 Id. at 1746.
52 Id.
53 Id. at 1748 (quotation marks and citation omitted).
54 Id.
55 Id. at 1748–49 (Ginsburg, J., dissenting).
56 Id. at 1750.
57 Id.
58 Id.
59 Id.
60 Id. at 1751 n.5.
61 Id. at 1751.
62 Id.
63 Id.
III. Two June 2018 Compelled-Speech Decisions

Although the *Masterpiece* majority did not decide the case on compelled-speech grounds, the Court resolved two other cases in June 2018—*NIFLA* and *Janus*—exclusively on that basis. Both of them shed light on how courts should analyze the speech question left undecided in *Masterpiece*.

A. *NIFLA* v. Becerra

The Court in *NIFLA*, ruling on a motion for preliminary injunction, held that a California statute mandating certain speech by pro-life pregnancy centers likely violates the First Amendment.64 The pregnancy centers that challenged that law come alongside women experiencing unexpected pregnancies, provide them with tangible resources and emotional support, and encourage them to keep their babies. The California law required medically licensed pro-life centers to “provide a government-drafted script about the availability” of state-funded abortions, “as well as contact information for how to obtain them.”65 The Court called this the “licensed notice.” California also mandated that the remaining pro-life centers—those that are not medically licensed—include in all their digital and print advertisements a 29-word disclaimer about their nonmedical status in multiple languages and font at least as large as the text of the advertisement itself.66 The Court referred to this as the “unlicensed notice.” The pregnancy centers argued to the Supreme Court that both of these requirements compelled them to speak unwanted messages in violation of their First Amendment rights.

The Court, in a 5–4 decision, held that the law likely violates the First Amendment. The majority opinion, written by Justice Thomas, began by announcing that the licensed notice “is a content-based regulation of speech.”67 Whenever a law “compel[s] individuals to speak a particular message,” it “alter[s] the content of [their] speech” and qualifies as a content-based regulation.68 Because California forced the licensed pro-life centers “to inform women how they can obtain state-subsidized abortions—at the same time [they] try to dissuade women from choosing that option—the licensed notice plainly ‘alters the content’ of [their] speech.”69

The Court refused to afford lesser constitutional protection to “a separate category of speech” that some “Courts of Appeals have [labeled as] ‘professional speech.’”70 “Professional speech” is speech by “individuals who provide personalized services to clients and who are subject to a . . . licensing and regulatory regime” when that speech “is based on their expert knowledge and judgment or . . . within the confines of the professional relationship.”71 The majority declined to carve out this subset of speech and subject it to lesser constitutional protection because it was concerned that doing so would enable governments to “suppress unpopular ideas”72 and deprive the people of “an uninhibited marketplace of ideas in which truth will ultimately prevail.”73 “States cannot choose the protection that speech receives under the First Amendment, as that would give them a powerful tool to impose ‘invidious discrimination of disfavored subjects.’”74 Nor did the majority accept California’s argument that laws compelling speech are problematic only if they “require a statement or endorsement of belief,”75 “hamper [the speaker’s] ability to present [its] own messages,”76 or are understood by viewers as the speaker’s “self-expression.”77 The Court was uninterested in tackling these requirements onto the compelled-speech doctrine.

After discussing the First Amendment flaws with the licensed notice, the majority found constitutional infirmities in the unlicensed notice. For that notice, California targeted “a curiously narrow subset of speakers,”78 raising the specter that it “has left unburdened those speakers whose messages are in accord with its own views.”79 And the majority said that the mandated notice “unduly burden[ed]” the pro-life centers’ protected speech80 by “drown[ing] out the facility’s own message” and “effectively rul[ing] out” many forms of advertising.81

Justice Kennedy wrote a concurring opinion. In it, he gave a particularly stinging rebuke to California and a ringing endorsement of freedom from compelled speech:

> The California Legislature included in its official history the congratulatory statement that the Act was part of California’s legacy of “forward thinking.” But it is not forward thinking to force individuals to be an instrument for fostering public adherence to an ideological point of view they find unacceptable. It is forward thinking to begin by reading the First Amendment as ratified in 1791; to understand the history of authoritarian government as the Founders then knew it; to confirm that history since then shows how

---

64 *NIFLA*, 138 S. Ct. at 2378.

65 Id. at 2371.

66 Id. at 2369–70.

67 Id. at 2371.


69 Id. (quoting Riley, 487 U.S. at 795). While laws that alter the content of speech ordinarily must survive strict scrutiny, the majority determined that it need not apply that standard because “the licensed notice cannot survive even intermediate scrutiny.” Id. at 2375.

70 Id. at 2371.

71 Id. at 2374 (quoting Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 641 (1994)).

72 Id. at 2374 (quoting *McCullen v. Coakley*, 134 S. Ct. 2518, 2529 (2014)).

73 Id. at 2375 (quoting *Cincinnati v. Discovery Network, Inc.*, 507 U.S. 410, 423–24 n.19 (1993)).


75 Id. at 42 (capitalization omitted).

76 Id. at 43.

77 *NIFLA*, 138 S. Ct. at 2377.

78 Id. at 2378 (quoting *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 580 (2011)).

79 Id. at 2377.

80 Id. at 2378 (citation omitted).
relentless authoritarian regimes are in their attempts to stifle free speech; and to carry those lessons onward as we seek to preserve and teach the necessity of freedom of speech for the generations to come. Governments must not be allowed to force persons to express a message contrary to their deepest convictions. Freedom of speech secures freedom of thought and belief. This law imperils those liberties.  

B. Janus v. AFSCME

The day after the Court ruled in NIFLA, it issued its decision in Janus, in which a five-Judge majority struck down an Illinois law requiring government employees to pay agency fees to unions. Because the money that the state forced employees to pay funded union speech to which some objects, the case was fundamentally about compelled speech. Justice Alito wrote he Court’s opinion—joined by Chief Justice Roberts and Justices Thomas, Kennedy, and Gorsuch—that overruled Abood v. Detroit Board of Education and invalidated the challenged Illinois law.  

The majority’s opinion decried the “damage” that compelled speech creates. Compelling individuals to mouth support for views they find objectionable violates [a] cardinal constitutional command, and in most contexts, such effort would be universally condemned.” The Court quoted Thomas Jefferson, who “denounced compelled support for beliefs as ‘sinful and tyrannical.’” The Court also recognized that government efforts to mandate speech strike against the dignity of the compelled speakers who are treated as mindless mouthpieces rather than free and independent thinkers. “When speech is compelled, individuals are coerced into betraying their convictions. Forcing free and independent individuals to endorse ideas they find objectionable is always demeaning.” For that reason, compelling speech is even worse than mandating silence, and “a law commanding ‘involuntary affirmation’ of objected-to beliefs would require ‘even more immediate and urgent grounds’ than a law demanding silence.” 

Janus rejected the argument that “the First Amendment was not originally understood to provide any protection for the free speech rights of public employees.” Taking away free speech protection for public employees would mean overturning decades of landmark precedent, so the Court refused to strip free-speech protection from an entire category of speakers just as it refused to do in NIFLA with respect to professionals.  

The Court also said that issues relating to “sexual orientation”—topics on which the Illinois unions speak—are “controversial subjects,” “sensitive political topics,” and “matters of profound value and concern to the public.” Speech on such vital and sensitive matters “occupies the highest rung of the hierarchy of First Amendment values,” and it “merits special protection.” With this decision in Janus, the Court ended its term with the free-speech splash that many expected.

What do NIFLA and Janus mean for the compelled-speech issue left unresolved in Masterpiece? The remainder of this article will explain how they support the compelled-speech arguments of creative professionals like Jack Phillips.

IV. Two Ongoing Cases Illustrating Two Open Questions

Two speech-related questions await answers in the wake of Masterpiece—one statutory, and one constitutional. The statutory question is whether public-accommodation laws are properly interpreted when they are used to punish creative professionals who decline requests for custom expression because of the messages that the requested expression will communicate rather than the protected status of the requester. And the constitutional issue is whether governments violate the First Amendment when they apply public-accommodation laws to require those professionals to create speech or art that expresses views in conflict with their conscience.

Lexington-Fayette Urban County Human Rights Commission v. Hands On Originals—a case pending before the Kentucky Supreme Court—illustrates the contexts in which courts are considering the statutory question. That case arose out of a complaint filed against Hands On Originals (HOO), a promotional print shop in Lexington, Kentucky, owned and managed by Blaine Adamson. Adamson and his shop serve all people, but they will not print all messages; as a matter of conscience, Adamson declines to create materials with messages that are contrary to his faith. In 2012, Adamson declined to print shirts promoting a gay pride festival. He did so because what he was asked to print—the words “Lexington Pride Festival” over a rainbow-colored logo—communicates messages about human sexuality that conflict with his religious beliefs. Wanting to help that customer as far as his conscience would allow, Adamson offered to connect him to another business that would print the shirts.

The local human-rights commission determined that Adamson violated the public-accommodation ordinance because his “objection to the printing of the t-shirts was inextricably intertwined with the status of the sexual orientation of members” of the group requesting the shirts. But that decision was reversed

82 Id. at 2379 (Kennedy, J., concurring) (citations, quotation marks, and alterations omitted).
84 Janus, 138 S. Ct. at 2460, 2486.
85 Id. at 2464.
86 Id. at 2463.
87 Id. at 2471.
88 Id. at 2464.
90 Id. at 2469.
91 Id.
92 Id. at 2476.
93 Id. (quoting Snyder v. Phelps, 562 U.S. 443, 452 (2011)).
94 Id. (quoting Snyder, 562 U.S. at 452).
by the state trial court, which affirmed Adamson's freedom not to print messages at odds with his faith. The trial court found that “[t]here is no evidence in this record that HOO or its owners refused to print the t-shirts in question based upon the sexual orientation of [the requesting group] or its members”; rather, they “declined to print the t-shirts in question because of the [it] MESSAGE.” Subsequently, the Kentucky Court of Appeals, in a 2-1 decision, upheld the trial court's ruling. “Nothing in the . . . ordinance,” the lead opinion concluded, prohibits Adamson from declining orders because of their “viewpoint or message.”

That is because "a message in support of a cause or belief . . . is a point of view and form of speech that could belong to any person, regardless of classification.” Thus, declining to print a message does not constitute unlawful discrimination based on a person's status. The Kentucky Supreme Court granted review, and that is where the case now sits. The statutory speech question—whether declining to create speech because of its message violates a public-accommodation law—is squarely before that court.

Another case—Telescope Media Group v. Lindsey—presents a good picture of the constitutional speech question. There, the Minnesota Department of Human Rights clearly announced that under its view of the state's public-accommodation law, businesses that engage in wedding-related work, including those who create art or speech, must help celebrate same-sex marriages if asked. The owners of Telescope Media Group, Carl and Angel Larsen, are filmmakers and devout Christians who believe that marriage is the union of a man and a woman. They cannot in good conscience use their skills as artists to celebrate a different view of marriage. Because they want to create films celebrating marriage, they are defending their freedom to do so without violating their faith. If they lose their case and the state enforces the law against them, they face penalties that include jail time.

A federal district court in Minnesota ruled against the Larsens. It did not deny that films are speech, but it nevertheless concluded that the state can apply its public-accommodation law to require the Larsens to create films celebrating same-sex weddings. While acknowledging that the law will have the effect of requiring the Larsens to create speech that they would not otherwise make, the court held that the statute is "content-neutral and not subject to strict scrutiny."

The Larsens appealed to the United States Court of Appeals for the Eighth Circuit, where the case is now pending.

V. Answering the Statutory Question

In cases like Hands On Originals, courts will need to decide the statutory question—whether creative professionals violate public-accommodation statutes when they decline orders because they disagree with the message they are asked to create. The Masterpiece opinions have a lot to say about that, particularly Justice Kagan's concurrence, Justice Gorsuch's concurrence, and Justice Ginsburg's dissent. Each of those opinions—the first joined by Justice Breyer, the next by Justice Alito, and the final by Justice Sotomayor—espouse the view that a creative professional does not violate a typical public-accommodation law by declining to create an expressive item with a message that it would not speak for anyone. Even the state affirmed this view in Masterpiece.

Justice Kagan wrote that it is not unlawful for business owners to decline a request for an expressive item that "they would not have made for any customer" because in doing so they treat the requester "in the same way they would have treated anyone else—just as [the public-accommodation law] requires." Stated differently, she acknowledged that business owners do "not engage in unlawful discrimination" when they "would not sell [a] requested [item] to anyone." Justice Ginsburg likewise recognized that businesses do not violate public-accommodation laws by "refus[ing] to make [an] item with [a] requested message for any customer" because people who request that message are "treated as any other customer would have been treated—no better, no worse." And Justice Gorsuch observed that creators of expression do not "intend[] to refuse service because of a customer's protected characteristic" if "they would not sell the requested [item] to anyone."

All of this directly supports the statutory arguments of people like Blaine Adamson in Hands On Originals. He declined the group's request for shirts promoting the pride festival because the message on them conflicted with his faith. He would not print shirts with that message for anyone, regardless of their protected status. He thus treated that group "as any other customer would have been treated—no better, no worse.” That does not violate public-accommodation laws. Justices Ginsburg, Sotomayor, Kagan, Breyer, Gorsuch, and Alito—either by authoring or

---

99 Id. at 13.
100 Id. at *7.
101 Id. at 1735 (Gorsuch, J., concurring).
102 Id. at 1733 n.*; see also id. (“A vendor can choose the products he sells, but not the customers he serves.”).
103 Id. at 1750 (Ginsburg, J., dissenting).
104 Id. at 1735 (Ginsburg, J., concurring).
105 Id. at 1750 (Ginsburg, J., dissenting).
106 Id. at 1735 (Gorsuch, J., concurring).
107 Id. at 1750 (Ginsburg, J., dissenting).
joining an opinion in *Masterpiece*—all recognized that. And none of the other Justices said anything in *Masterpiece* suggesting that they disagree with that view.

This understanding of public-accommodation laws makes perfect sense because a contrary reading would conflict with the purpose of such laws, which is to require equal treatment. Creative professionals like Adamson already treat people equally—if they can in good conscience print a specific message, they will do so for anyone who requests it. What their opponents demand is that they create speech that they will not make for anyone, effectively entitling particular customers to preferred—which is not equal—treatment and forcing creative professionals like Adamson to expand the services they offer. No reasonable interpretation of public-accommodation laws requires that.

These principles, which are straightforward in a case like *Hands On Originals*, become more difficult to apply in the wedding context. *Masterpiece* illustrates this. On the one hand is the view of Justices Kagan and Ginsburg. They said that the relevant message in *Masterpiece* was a generic expression of celebration for a wedding. On the other hand is Justice Gorsuch’s approach. He said that the requested message was more specific—an expression of celebration for a specific kind of wedding. Courts struggling to decide the statutory question in wedding cases will vacillate between those two approaches and maybe even invent other theories along the way. But regardless of where courts land on that issue, it seems likely that sooner or later some of them will need to face the constitutional speech question left unresolved in *Masterpiece*.

VI. Answering the Constitutional Question

On the constitutional speech issue, *Masterpiece*, *NIFLA*, and *Janus* together support the argument that creative professionals who believe that marriage is the union of a man and a woman should be able to live consistently with that belief. Not only are such beliefs “decent and honorable” as the Court noted when announcing that the Constitution requires states to recognize same-sex marriages, but people must be free to “car[ry]” those beliefs “into the public sphere or commercial domain.”

A. No Classes of Speakers Should be Excluded from First Amendment Protection

*NIFLA* and *Janus* cast grave doubt on the primary speech argument that the state raised in *Masterpiece*. The state argued that the compelled-speech doctrine offers no protection when governments apply public-accommodation laws to people who earn a living creating and selling speech. But *Janus* expressly rejected a similar categorical argument that sought to exclude a class of speakers from the First Amendment’s protection against compelled speech. The unions claimed that the First Amendment does not protect the free-speech rights of public employees, but *Janus* found no merit to that claim. *NIFLA* further explained why the Court is skeptical of attempts to strip some speakers of full First Amendment protection. Accepting the state’s argument that professional speech is subject to reduced constitutional scrutiny, the Court explained, would empower states to manipulate public discussion and deny citizens an “uninhibited marketplace of ideas in which truth will ultimately prevail.” The best test of truth is the power of the thought to get itself accepted in the competition of the market, and the people lose when the government is the one deciding which ideas should prevail.

Similar harm would result if governments were free to wield public-accommodation laws to compel speech without stringent First Amendment oversight. They could, for example, single out speakers whose views they do not like by targeting them for punishment when they decline to speak messages the government favors. By banishing one side of a deeply divisive issue, the government robs the people of the “uninhibited marketplace of ideas” that the First Amendment promises and distorts the search for “truth.” This is especially dangerous where the government treats people with unpopular views worse than those with popular views. The five Justices who joined the majority opinions in *Masterpiece*, *NIFLA*, and *Janus* resisted efforts to strip First Amendment protection from categories of speakers. Those arguments will likely face a similar fate in a future case where the government applies a public-accommodation law to compel speech.

B. No Extra Elements Should be Added to a Compelled-Speech Claim

Governments seeking to enforce public-accommodation laws in cases like *Masterpiece*, *Hands On Originals*, and *Telescope Media Group* also argue that simply proving compelled speech is not enough for a compelled-speech claim. They attempt to engraft other requirements onto the compelled-speech doctrine, such as the need to show a burden on the compelled speaker’s expression or some other form of harm. The state advanced a similar argument in *NIFLA*, contending that the speech-compelling law posed no First Amendment concern because it did not “hamper [the speaker’s] ability to present [its] own messages.” The Court did not accept that argument, reasoning instead that courts, “as a general matter,” apply strict scrutiny when government officials


110 *Masterpiece*, 138 S. Ct. at 1729.


112 *Janus*, 138 S. Ct. at 2469–70.

113 Id.

114 Id.

115 *NIFLA*, 138 S. Ct. at 2371–72, 2374–75.

116 Id. at 2374 (quoting *McCullen*, 134 S. Ct. at 2529).

117 Id. at 2375 (alterations, quotation marks, and citation omitted; emphasis added).

118 Id. at 2374.

119 E.g., *Craig*, 370 P3d at 288 (concluding that Phillips’s speech was not burdened by forcing him to create custom art celebrating same-sex marriage because he was still free to “express[] [his] views on same-sex marriage” and “to disassociate [him]self from [his] customers’ viewpoints”).

force individuals to “alter the content of [their] speech.”121 Thus, under *NIFLA*’s logic, when government officials use public-accommodation laws to force individuals to create expression, that alters the content of their speech, and strict scrutiny applies.122

While *NIFLA* shows that compelled speakers need not demonstrate a burden on their speech, *Janus* tells us that the harm to coerced speakers is inherent and intolerable. When the government mandates that people speak against their consciences, it forces them to “betray[] their convictions” and “endorse ideas they find objectionable.”123 That “is always demeaning,”124 and the First Amendment demands “immediate and urgent grounds” to justify such a deep intrusion into conscience.125 Since, as *Janus* says, forcing people “to subsidize . . . speech” that they oppose is “tyrannical,” compelling them to create and then distribute such speech—which is what the government requires in cases like *Masterpiece, Telescope Media Group,* and *Hands On Originals*—should be unthinkable.126 When governments do that, their actions should be, in *Janus*’s words, “universally condemned.”127

In addition, governments routinely insist that the existence of a compelled-speech violation depends on the perceptions of viewers.128 But the Court in *NIFLA* declined to use third-party perceptions to excuse compelled speech. It considered only whether the law forced the parties to “alter[] the content” of their speech, paying no mind to what others might think about who was speaking.129 Similarly, Justice Thomas’s *Masterpiece* concurrence rejected the Colorado Court of Appeals’ conclusion that Phillips was not entitled to compelled-speech protection because “a reasonable observer would think he is merely complying with Colorado’s public-accommodations law.”130 The “Court has never accepted” that argument, Justice Thomas explained, because to do so “would justify any law that compelled protected speech.”131 Moreover, jettisoning reliance on viewers’ perceptions accords with *Janus*’s recognition that “[f]orcing free and independent individuals to endorse ideas they find objectionable is always demeaning,”132 regardless of what others might think.

### C. The Constitution Protects the Decision to Decline a Request Because of the Message It Would Communicate

In contrast to governments’ attempts to diminish the protection provided by the compelled-speech doctrine, Phillips and others who make speech or art for a living argue that compelled-speech principles protect them in a narrow set of circumstances: when a customer asks them to create expression, and they decline the request because of the message communicated by the item rather than the protected characteristic of the requesting person. The vision that the *Masterpiece* majority casts for how courts should decide the compelled-speech issue in future cases supports this constitutional protection, as does Supreme Court precedent.

The *Masterpiece* majority was clearly open to protecting individuals with “religious and philosophical objections to gay marriage” where they are engaged in “protected forms of expression.”133 But, the Court counseled, “any decision in favor of the baker would have to be sufficiently constrained.”134 Constitutional protection limited to creative professionals would square with this vision, the Court implicitly recognized, because “there are no doubt innumerable goods and services that no one could argue implicate the First Amendment.”135 Among creative professionals, the Court appeared to draw the same line that Phillips embraced during the litigation. On the one hand are artists and creators of expression who “refuse[] to sell” even nonexpressive items “for gay weddings.”136 The Court was not disposed to afford them constitutional protection.137 On the other hand are those who, like Phillips, serve everyone and sell their ready-to-purchase goods to anyone, but who decline “to use [their] artistic skills to make an expressive statement, a wedding endorsement in [their] own voice and of [their] own creation” in violation of their convictions.138 In those instances, the customers’ request constitutes “a demand for [the creative professionals] to exercise the right of [their] own personal expression” in conflict

---

121 *NIFLA*, 138 S. Ct. at 2371.

122 *NIFLA* also reaffirmed that the Supreme Court does not look favorably on efforts to subject content-based speech regulations to lesser constitutional scrutiny, stating that it “has been especially reluctant to exempt a category of speech from the normal prohibition on content-based restrictions.” 138 S. Ct. at 2372 (citation omitted).

123 *Janus*, 138 S. Ct. at 2464. And as Justice Kennedy affirmed in his *NIFLA* concurrence, allowing the government to compel speech also “imperils” “freedom of thought and belief.” 138 S. Ct. at 2379.

124 *Janus*, 138 S. Ct. at 2464.

125 *Id.* (quoting *Barnette*, 319 U.S. at 633).

126 *Id.* at 2463–64.

127 *Id.* at 2463.

128 E.g., *Craig*, 370 P.3d at 286 (holding that Phillips “does not convey a message supporting same-sex marriages merely by abiding by the law” because “a reasonable observer would understand that [his] compliance with the law is not a reflection of [his] own beliefs”); Br. for State Resp’ts at 43, *NIFLA* v. Becerra, No. 16-1140 (U.S. Feb. 20, 2018) (arguing that no compelled-speech violation occurred because no reasonable viewer would understand the speech to be the compelled speaker’s own “self-expression”).

129 See *NIFLA*, 138 S. Ct. at 2371–76 (ignoring third-party perceptions in the compelled-speech analysis).

130 *Masterpiece*, 138 S. Ct. at 1744 (Thomas, J., concurring).

131 *Id.*

132 *Janus*, 138 S. Ct. at 2464.

133 *Masterpiece*, 138 S. Ct. at 1727.

134 *Id.* at 1728 (“Any decision in favor of the baker would have to be sufficiently constrained.”).

135 *Id.*

136 *Id.*

137 *Id.* (recognizing that such a case “would be a different matter and the State would have a strong case”).

138 *Id.*
with their religious beliefs. The Court seemed willing to apply First Amendment protection under those circumstances.

Doing so would be consistent with Supreme Court precedent. In *Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston*, the Court distinguished between a public accommodation’s objection to a requested message and a decision to turn people away because of who they are. In that case, parade organizers allowed LGBT individuals to participate in their parade, but they declined a request for an LGBT group to march behind its own pro-LGBT banner. The organizers did not “exclude the [group’s members] because of their sexual orientations,” but because of the messages that they wanted to communicate by “march[ing] behind [their] banner.” Similarly in *Masterpiece, Hands On Originals, and Telescope Media Group*, the business owners serve LGBT individuals; they simply decline to create art or expression that celebrates views that violate their faith. As the Court did in *Hurley*, courts should incorporate into their compelled-speech analysis the distinction between refusing to serve people because of who they are and declining to express a message because of a disagreement with it.

**D. Responding to Counterarguments**

Some contend, like the state did in *Masterpiece*, that it is not workable for courts to determine whether a business owner’s custom work qualifies as speech. NIFLA implies that such an argument will not gain much traction: “While drawing the line between speech and conduct can be difficult, this Court’s precedents have long drawn it, and the line is long familiar to the bar.” In light of this, it is unlikely that the Court will agree that courts either cannot or should not distinguish between speech and nonspeech. In fact, the Court has frequently drawn that very line by asking whether a particular item “communicate[s] ideas” and whether it is analogous to other kinds of protected speech.

Others argue that protecting creative professionals’ freedom from compelled speech could harm the dignity of would-be customers. But their arguments ignore all the ways in which the state is free to shield the dignity of consumers. Where speech is not involved, which *Masterpiece* recognized is the case with “innumerable goods and services,” the compelled-speech doctrine provides no protection, and customer requests cannot be declined on that basis. This means that even if someone creates and sells speech for a living, the compelled-speech doctrine does not apply when they sell items that are not speech. So in *Masterpiece*, while Phillips can decline requests to design custom artistic wedding cakes to celebrate same-sex weddings, he cannot refuse to sell his generic, non-expressive brownies or cookies to anyone. And even when artists like Phillips create expression like a custom-painted cake with words and images, once they finish creating that expression and offer it for sale, they cannot decline to sell it to anyone. Compelled-speech protection is thus narrow enough that it may be afforded without inflicting “a community-wide stigma” on, or a “serious diminishment to [the] dignity and worth” of, any particular group.

To be sure, in the limited situations where the compelled-speech doctrine affords protection, a customer still might allege a dignitary harm. Yet in that narrow circumstance, dignitary interests are at stake for the compelled speaker too. As *Janus* said, it “is always demeaning” to compel people to speak messages that “betray[] their convictions” or “endorse ideas they find objectionable.” This is consistent with the Court’s longstanding recognition that free-speech protection safeguards “individual dignity” and that a third-party’s offense at a decision not to speak is not a sufficient reason to coerce expression. Justice Thomas echoed the latter point in his *Masterpiece* concurrence, explaining that dignity-based “justifications” for compelled speech are “completely foreign to [the Court’s] free-speech jurisprudence.” States cannot punish protected speech because some group finds it offensive, hurtful, stigmatic, unreasonable, or undignified.

139 *Id.*


141 *Id.* at 572 (distinguishing an “intent to exclude homosexuals” from a “disagreement” with a message).

142 *Id.*


144 Some claim that the Supreme Court rejected this distinction in *Christian Legal Society v. Martinez*, 561 U.S. 661, 689 (2010). But while the Supreme Court rejected a status/conduct distinction in *Martinez*, see *Id.* (“Our decisions have declined to distinguish between status and conduct” in the sexual-orientation context), it recognized a status/message distinction in *Hurley*. Creative professionals like Jack Phillips do not differentiate between their customers’ status or conduct—but, for example, serving LGBT customers who are not in same-sex relationships but refusing to serve those who are. Instead, like the parade organizers in *Hurley*, the decisionmaking of people like Jack Phillips does not depend on a customer’s status or conduct, but only on the messages that customers ask them to express through their custom artistic creations.

145 NIFLA, 138 S. Ct. at 2373 (quotation marks and citations omitted).


147 *Masterpiece*, 138 S. Ct. at 1728.

148 *Id.* at 1727.

149 *Janus*, 138 S. Ct. at 2464.


151 *Hurley*, 515 U.S. at 574 (declining to compel speech because “the point of all speech protection . . . is to shield just those choices of content that in someone’s eyes are . . . hurtful.”); see also Matal v. Tam, 137 S. Ct. 1744, 1764 (2017) (plurality opinion) (rejecting an asserted “interest in preventing speech expressing ideas that offend” because “we protect the freedom to express the thought that we hate”) (quotation marks omitted); *Snyder*, 562 U.S. at 458 (“If there is a bedrock principle underlying the First Amendment, it is that the government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.”) (quoting Texas v. Johnson, 491 U.S. 397, 414 (1989)); *Boos v. Barry*, 485 U.S. 312, 322 (1988) (expressing grave doubts about the government’s “interest in protecting the dignity” of listeners from harmful speech since that is “inconsistent with our longstanding refusal to punish speech because the speech in question may have an adverse emotional impact on the audience”) (quotation marks, alterations, and citation omitted).

152 *Masterpiece*, 138 S. Ct. at 1746 (Thomas, J., concurring).

153 *Id.*
VII. Conclusion

Masterpiece, NIFLA, and Janus together urge future courts to uphold the expressive freedom of creative professionals who serve all people but decline to create speech that communicates messages in violation of their consciences. This narrow constitutional protection honors the directive in Masterpiece to respect the religious beliefs and fundamental freedoms of people who create speech for a living while simultaneously respecting the dignity of customers. It guarantees that governments do not “demean[]” creative professionals by forcing them to say what is not in their mind,154 while also ensuring that no group of customers experiences a “serious diminishment to their own dignity and worth.”155 It provides a promising path forward and a reasonable resolution to a contentious national debate.

154 Janus, 138 S. Ct. at 2464.

155 Masterpiece, 138 S. Ct. at 1727.
Intellectual Property

THE SUPREME COURT TACKLES PATENT REFORM:
A SERIES OF ARTICLES EXAMINING OIL STATES ENERGY SERVICES, LLC v. GREENE’S ENERGY GROUP, LLC

By Richard A. Epstein

Note from the Editor:
This article argues that the Supreme Court should find unconstitutional the America Invents Act’s inter partes review procedure for administratively reviewing patents. It is the first in a series of articles by Professor Epstein about the issues at stake in Oil States Energy Services, LLC v. Greene’s Energy Group, LLC. Earlier versions of these articles were posted on the Fed Soc Blog.

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.


About the Author:
Richard A. Epstein is the Laurence A. Tisch Professor of Law at NYU School of Law, the Peter and Kirsten Bedford Senior Fellow at the Hoover Institution, and the James Parker Hall Distinguished Service Professor of Law Emeritus and Senior Lecturer at the University of Chicago.

PRE-ARGUMENT ARTICLE:
WHY THE SUPREME COURT SHOULD END INTER PARTES REVIEW IN OIL STATES

Oil States Energy Services, LLC v. Greene’s Energy Group, LLC is the most important intellectual property case to come before the Supreme Court in many years.1 It challenges some of the innovative dispute resolution provisions of the 2011 American Invents Act (AIA), the most significant legislative reform of patent law since the Patent Act of 1952.2 Oil States assumes its vast significance because its outcome will determine, perhaps for decades, the litigation framework for all future patent disputes. Although widely hailed as a statute that aids inventors, the AIA has been subject to searching criticism that it amounts to “Dubious Patent Reform”3 driven by “The Myth of Patent Quality.”4 The case for the far-reaching reforms of the AIA rests on the common claim, made in the legislative history, that weak patent claims had routinely been approved under the earlier patent regime. But even if weak patents were a problem before the AIA, the 2011 legislation applies a sledge hammer where only a scalpel was warranted. This ham-handed response did more than threaten weak patents; it also undercut the safety of strong patents, since the procedure it instituted—inter partes review (IPR)—offers greater advantage to accused patent infringers than the traditional litigation process in district court. For this reason, parties accused of infringement have beaten a steady path to the Patent Trial and Appeal Board (Board or PTAB), where these cases are tried inside the U.S. Patent & Trademark Office (PTO).5

The contrast between the 1952 Act and the AIA could not be starker. The 1952 Act was drafted by an eminent committee

of two patent experts: Giles Rich,6 who later went on to serve 43 years as a patent appeals judge, and Pasquale Joseph Federico, who served as a high ranking official in the PTO. It was passed by voice vote. The 1952 Act is less than half the length of the AIA, but it is more than twice as good. The AIA is cluttered with arcane refinements and complex procedures, and it violates the fundamental maxim of legislative reform: if it ain't broke, don't fix it. The AIA is in sore need, as former Chief Judge Paul Michel recently argued,7 of urgent substantive repair.

At issue in Oil States is the constitutionality of IPR which, as the PTO describes it:

[I]s a trial proceeding conducted at the Board to review the patentability of one or more claims in a patent only on a ground that could be raised under §§ 102 [dealing with novelty and prior art] or 103 [dealing with nonobvious subject matter], and only on the basis of prior art consisting of patents or printed publications.8

These are elaborate trial proceedings before panels of indefinite size appointed by the head of the PTO. The panels render final judgments in IPR cases, to which appeal may only be had to the Federal Circuit, which like other appellate bodies does not supply de novo review on all disputed issues. Contrasting the PTAB procedures with those of the Trademark Trial and Appeal Board (TTAB) highlights how far the former fall short of satisfying the requirements of separation of powers and due process. The rules in IPR proceedings are different from those which apply in disputes before the TTAB, where, as Justice Alito noted in B & B Hardware v. Hargis Industries:

[A]fter the TTAB decides whether to register the mark, a party can seek review in the U.S. Court of Appeals for the Federal Circuit, or it can file a new action in district court. See 15 U.S.C. § 1071. In district court, the parties can conduct additional discovery and the judge resolves registration de novo.9

Parties who challenge TTAB decisions may bring their disputes before federal trial courts, thereby benefitting from the separation of powers and preserving the procedural protections guaranteed to them by the Constitution.

In Oil States, IPR took place simultaneously with an infringement action that Oil States brought against Greene Energy in federal district court. Nearly one year after Oil States sued Greene Energy in federal district court for patent infringement, Greene petitioned the PTAB for IPR, which was granted. The federal district court then construed the disputed patent in ways that resolved the claim conclusively in favor of Oil States. But in the IPR proceeding, the Board upheld the challenge to the patent on the grounds that its claims had been anticipated in prior art. The Board refused to allow Oil States to amend its claims, and Oil States appealed to the Federal Circuit, which affirmed the Board’s decision without opinion. The case is now before the Supreme Court to resolve the question of whether a decision made inside the PTO can displace that of a federal district court.

The issues raised in Oil States do not put the spotlight on the AIA’s substantive problems. Instead, the case calls into question the procedures that the AIA uses to examine, and then reexamine, the validity of patents that are challenged on two grounds frequently at play in patent disputes. Under the AIA, the accused patent infringer has the right to remove a case from a district court to a specially constituted panel inside the PTO. Greene’s Energy Group, the respondent in Oil States, requested that the PTO conduct an IPR of an Oil States patent dealing with drilling equipment. That request was granted, and the PTAB established a special panel to review the case; the PTO panel determined that some of Oil States’ claims were not patentable. The members of that panel were individually appointed for this specific case by Michelle K. Lee, then director of the PTO, from a roster of some 200 potential judges. Under the AIA, PTO panels need not be constructed on a once-and-for-all basis, for the PTO reserves the power to appoint additional judges to any panel in order to “secure and maintain uniformity of the Board’s decisions.”10

Notwithstanding the substantive importance of the resolution of these disputes, once the case is transferred to the PTO, the patent holder no longer enjoys the procedural protections that normally attach to litigation in an Article III federal court. These protections include a trial before an independent judge with lifetime tenure and the constitutional right to a jury trial that is explicitly protected under the Seventh Amendment.

By challenging this scheme, Oil States raises critical separation of powers and due process questions. Separation of powers is not a merely academic concern. As the Supreme Court has repeatedly held, Framers of the Constitution separated and divided the powers of government in order to diffuse power, thereby to better secure liberty.11 They set forth the powers of the legislative branch in Article I, the executive in Article II, and the judicial in Article III. Whether patent adjudication is properly the province of legislative or Article III courts—tribunals created by an act of Congress like the AIA, as opposed to courts established under Article III of the Constitution itself—depends on an arcane question: Should patents be classified as “public rights” over which the government has extensive discretion in setting the rules for...

---


their validity and enforcement, or as private property, creating a set of vested rights, enforceable at common law and in courts of equity, that are not subject to the same overall level of legislative manipulation.

The Solicitor General claims in his brief for the federal respondent that a patent “is a paradigmatic public right,” which gives the government the right not only to grant patents, but to remove them or to grant patents, but to remove them as well by actions solely within the executive branch through the PTO. The brief further claims that “the traditional understanding of patent rights [is that they are] privileges that the government may revoke without judicial involvement,” and, apparently without any compensation at all. On this view, whenever the existence of a right depends on the will of Congress, Congress can set the terms for how that right will be adjudicated. The implications of this claim are breathtaking, because it means that whenever Congress creates any statutory rights, it can extend Article III courts of their jurisdiction. The government position thus inverts the basic rule of separation of powers, under which Congress decides the content of new federal causes of action, but cannot remove them from federal court. Article III, Section 2 under their authority.” To that command, the government adds the caveat: “unless Congress decides otherwise,” which makes Article III an optional extra. The brief submitted by Greene Energy stresses these same themes. A careful history of the public rights doctrine shows that it is far more limited in its scope than the government and Greene Energy claim; it is chiefly limited to incidental administrative disputes that deal with the disposition of government property and the settling of financial disputes in the administrative scheme.

I. Private Property Versus Public Rights

A. Origin of the Public Rights Exception

The first major statute dealing with patents was the English Statute of Monopolies, adopted in 1623, which made it clear that “the force and validity of [patents], and every of them, ought to be, and shall be for ever hereafter examined, heard, tried, and determined, by and according to the common laws of this realm, and not otherwise.” As described by one commentator, “In this regard, the Statute of Monopolies was more than simply a restatement of existing law: it introduced a crucial change by granting jurisdiction to the common law courts in place of the monarch’s ‘act of grace.’”

It was this concern that helped shape the procedures by which disputes were to be resolved once patents were granted.

17  59 U.S. 272 (1855).
Fifth Amendment precisely because it was directed against “public debtors.” The Court then concluded that government actions for the reconciliation of accounts do not count as “a judicial controversy” that requires an adjudication by an Article III court. Justice Curtis wrote, “The power to collect and disburse revenue, and to make all laws which shall be necessary and proper for carrying that power into effect, includes all known and appropriate means of effectually collecting and disbursing that revenue, unless some such means should be forbidden in some other part of the constitution.” In short, although due process of law required that most legal disputes be adjudicated by courts, disputes involving “public rights” could be dealt with in administrative proceedings. The Court narrowly defined what counts as a public right. The full passage makes it clear that the public rights exception has nothing to do with disputes between private parties, including ordinary patent infringement cases, that necessarily fall exclusively within the scope of the judicial power:

To avoid misconstruction upon so grave a subject, we think it proper to state that we do not consider Congress can either withdraw from judicial cognizance any matter which, from its nature, is the subject of a suit at the common law, or in equity, or admiralty; nor, on the other hand, can it bring under the judicial power a matter which, from its nature, is not a subject for judicial determination. At the same time there are matters, involving public rights, which may be presented in such form that the judicial power is capable of acting on them, and which are susceptible of judicial determination, but which Congress may or may not bring within the cognizance of the courts of the United States, as it may deem proper. Equitable claims to land by the inhabitants of ceded territories form a striking instance of such a class of cases; and as it depends upon the will of congress whether a remedy in the courts shall be allowed at all, in such cases, they may regulate it and prescribe such rules of determination as they may think just and needful.

The customs dispute in Murray’s Lessee represents one class of case for which Congress may opt to resolve these claims in either a judicial or administrative proceeding. Disputes over government land grants in ceded territories, as from Mexico, are yet another. But all disputes that take place at common law, equity, and admiralty—the three explicit heads of judicial power under Article III—fall outside of the scope of this classification, and they must be adjudicated in Article III court proceedings.

Ordinary disputes in patent cases, where the remedies sought are damages at law, injunctions in equity, or both, fall squarely within the scope of the judicial power. But under the AIA, the phrase “public rights” has been ripped out of its original context such that Murray’s Lessee is read to justify treating ordinary patent disputes as administrative matters; this is an extremely broad, and almost certainly incorrect, reading of the public rights exception the case established.

B. The Supreme Court Has Rejected a Broad Reading of the Public Rights Exception

The Supreme Court unanimously rejected a broad reading of the public rights exception in the trademark context in its 2017 decision in Matal v. Tam.18 The PTO had sought to deny trademark registration to a band named “The Slants” under the so-called disparagement clause of the Lanham Act.19 In essence, without using the exact words, the government sought to portray the dispute as one involving public rights, and therefore amenable to adjudication within the PTO and not requiring a hearing in an Article III court.

One key issue in Tam was whether issuing a trademark is a form of government speech, as the PTO claimed, given that federal registration functionally confers a subsidy on the holder of the trademark. The theory is odd in all respects. Parties have to pay in order to receive these rights, which they get, not as a matter of political largesse, but as a matter of right, once it is established that their applications meet the requirements of substantive law. When the government protects intellectual property, it is not transferring some preexisting patented technology, copyrightable material, or trademark to a private party. It is instead recognizing that the private party is entitled to the exclusive right in its invention; such rules of first possession have allowed for the creation of private rights in an invention since Roman times. The government is correct to say that there are no patent rights in the state of nature, but the same theory of natural rights that allows for the acquisition of title by acquisition in land works just as well to allow for the acquisition of patent rights by invention of the patented item. Patent laws, grounded in the Constitution’s phrase, “to promote science and the useful arts,” are intended to ensure, as Adam Mossoff demonstrated, that Congress does not issue patents to any favored applicant as a matter of political largesse.20

The government’s treatment of patents thus differs totally from the conveyance of government lands. In the latter case, it transforms public property into private property. But government protection of intellectual property begins with an application by a private party that identifies the subject matter of the right that its own labors have created. The government system is intended to secure inventors’ and authors’ rights to the results of their intellectual labors, not to retain the power to dispense these rights as if the government had created the inventions itself.

Therefore, it is incorrect to treat the validation of intellectual property as a form of government speech, comparable to statements made by public officials on matters of public importance. Keeping the boundary lines clear is critical, for, as Justice Alito pointed out, treating intellectual property protection as government speech “would either eliminate any First Amendment protection [for private parties seeking IP protection] or result in highly permissive rational-basis review.” Given that possibility, the Court decisively rejected the view that trademark protection is a government subsidy, properly limiting that description to government transfer programs rather than extending it to systems used to protect common law rights. Accordingly, Justice Alito noted that trademark registration simply strengthened the common law

protection for trademarks, first by giving constructive notice to the rest of the world of the existence of the trademark, and second by creating a presumption of its validity.

Justice Alito’s opinion then reverted to a key theme in *Murray’s Lessee*: “Trademarks and their precursors have ancient origins, and trademarks were protected at common law and in equity at the time of the founding of our country.” He elaborated, showing how government-registration schemes are not subsidies: “For example, the Federal Government registers copyrights and patents. State governments and their subdivisions register the title to real property and security interests; they issue driver’s licenses, motor vehicle registrations, and hunting, fishing, and boating licenses or permits.” The point of these registration schemes is not to transfer wealth from one group to another, as would be the case if they were subsidies. It is to strengthen the system of property rights in ways that allow their owners to transfer, mortgage, and license their rights securely.

Patents fall within the category of properties that receive protection both at common law and in equity, which means that their validity can only be adjudicated in Article III courts. The fact that the Takings Clause of the Fifth Amendment has always protected patents—subject to some constraints—makes the patentee an exclusive property in the patented invention which cannot be appropriated or used by the government itself, without just compensation, any more than it can appropriate or use without compensation land which has been patented to a private purchaser.21 *Horne* follows an unbroken line of Supreme Court cases that takes exactly the same view.22 So far as the rights vested in the patent holder are concerned, there has not been an iota of movement between *Murray’s Lessee* and *Tam*.

C. Lower Courts Continue to Treat Patents as Public Rights

In dealing with these issues, however, the lower federal courts have deviated from the applicable Supreme Court precedents in a manner that obscures all the basic issues. There is no appellate opinion in *Oil States*, but the critical case on which the government relies for its broad account of public rights is the Federal Circuit’s 2015 decision in *MCM Portfolio LLC v Hewlett-Packard*.23 In that case, decided before *Tam*, Judge Timothy Dyk, writing for a unanimous panel, held that the public rights doctrine covered patent rights. This made it constitutionally permissible to use IPR before the PTO without running afoul of either the case or controversy limitation in Article III or the Seventh Amendment guarantee of a jury trial which applies to “suits at common law.”

*MCM* could not possibly stand under *Murray’s Lessee* on its own, so it is important to examine the precedents that Judge Dyk relied upon in the case24 to expand the class of “public rights” so that patent rights supposedly fall squarely within it. Yet a close examination reveals that none of these cases made any change in the basic pattern established in *Murray’s Lessee*. For example, in the 1921 case of *Block v. Hirsh*, a narrow five-to-four majority upheld a two-year rent control statute enacted in Washington, D.C. in the aftermath of World War I, when a huge influx of new residents put upward pressure on local rents.25 That D.C. statute also authorized the formation of an administrative board to settle disputes over what rent was reasonable under the ordinance. Justice Holmes held that, because the rent control statute had suspended the landlord’s common law right to regain possession of his property, it concomitantly allowed the government to adopt a collateral administrative process to determine whether the rent was reasonable under a set of government standards, not the private contract. Given the circumstances of the case, the source of that right was the statutory command, not the common law. Yet the government’s brief in *Oil States* overstates the scope of the public rights exception applied in *Block* when it states that it “upheld resolution of landlord tenant disputes through a federal administrative system”; in reality, all other landlord-tenant disputes on such matters as breach of warranty were outside its scope of the system approved in *Block*. Furthermore, even while upholding Congress’ procedure, Holmes noted that “on [the reasonableness of the rents imposed] question the courts are given the last word.” This is hardly the case in the IPR system.

Judge Dyk also cited a 1929 Supreme Court case, *Ex Parte Bakelite Corporation*, which examined the validity of a Tariff Commission constituted “to protect domestic industry and trade against ‘unfair methods of competition and unfair acts’ in the importation of articles into the United States.”26 Justice Van Devanter noted that the Tariff Commission and similar bodies were “legislative courts” that Congress could create outside the scope of Article III. He pointed out that “[c]onspicuous among such matters”—where a private party challenges a tariff determination made by the Tariff Commission, or the rent set by the rental board as in *Block*—“are claims against the United States.” The category of public rights also includes other claims against the United States, including both the resolution of private claims in lands ceded to the United States in the territories and specialized courts for resolving tribal disputes. But that category “include[s] nothing which inherently or necessarily requires judicial determination.” Nowhere in *Bakelite* is it argued or even hinted that the class of public rights could include actions at common law or in equity against private defendants. It plainly does not. Furthermore,

---

24 It is worth noting that *Tam* had not yet been decided when Judge Dyk wrote his decision in *MCM*.
one main difference between the legislative courts considered in *Bakelite* and Article III courts is that legislative courts may be empowered to give advisory opinions, while Article III limits courts to deciding only cases and controversies. Judge Dyk did not refer to the narrow constitutional place of legislative courts, on which the *Bakelite* Court relied, in his one-sentence analysis of the case. The government’s cryptic discussion of the issue in its *Oil States* brief wholly glosses over the role of Article I courts.

Judge Dyk next discussed the 1985 case of *Thomas v. Union Carbide Agricultural Products Co.*27 That case upheld a complex procedure under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). The statute allowed the EPA to use data submitted by a pesticide manufacturer when considering an application by a rival firm only if that rival firm agreed to binding arbitration, with limited judicial review, if it failed to reach a compensation agreement with the submitting firm. *Thomas* involved the resolution of technical issues within the administrative agency itself, which is an uncontroversial use of administrative proceedings. The Court took the additional small step to permit the agency to turn the dispute over to binding arbitration for convenience, but it did not allow these informal, non-judicial procedures to resolve any subsequent dispute over the validity of the registration of the various substances covered by the Act, a point that was not noted in the government’s brief in *Oil States*. *Thomas* is therefore similar to *Bakelite*, which also explicitly allowed for the creation of legislative courts to ease the burden of what would otherwise be purely administrative systems. But, also like *Bakelite*, it does not even hint at the possibility that similar procedures could be used—in place of litigation in federal court—to conclusively resolve disputes between private parties over the validity of the registration in which damages and injunctions were sought.

The last of the cases that Judge Dyk addressed is the well-known 2011 bankruptcy decision of the Supreme Court, *Stern v. Marshall*, which repeatedly stressed the need to resolve ordinary common law disputes in Article III courts.28 At issue in *Stern* was whether the plaintiff, as executor for the estate of Vicki Marshall, could bring in bankruptcy court—an Article I legislative court—a common law tort claim against the executor of E. Pierce Marshall’s estate.29 Chief Justice Roberts emphatically rejected the possibility:

The Bankruptcy Court in this case exercised the judicial power of the United States by entering final judgment on a common law tort claim, even though the judges of such courts enjoy neither tenure during good behavior nor salary protection. We conclude that, although the Bankruptcy Court had the statutory authority to enter judgment on Vickie’s counterclaim, it lacked the constitutional authority to do so . . . .28

. . . Under “the basic concept of separation of powers . . . that flow[s] from the scheme of a tripartite government” adopted in the Constitution, “the ‘Judicial Power of the United States’ . . . can no more be shared” with another branch than “the Chief Executive, for example, can share with the Judiciary the veto power, or the Congress share with the Judiciary the power to override a Presidential veto.”

Nonetheless, Judge Dyk treated the Supreme Court’s rejection of an expanded role for the bankruptcy courts to cover common law tort claims as though it supports an expanded role for the PTO in evaluating common law patent claims. Despite the outcome in *Stern*, he accepted the proposition that IPR under the AIA falls within the class of public rights exceptions to the requirements of Article III. He thus wrote for the panel:

The patent right “derives from an extensive federal regulatory scheme,” *Stern*, 131 S.Ct. at 2613, and is created by federal law. Congress created the PTO, “an executive agency with specific authority and expertise” in the patent law, and saw powerful reasons to utilize the expertise of the PTO for an important public purpose—to correct the agency’s own errors in issuing patents in the first place. . . . There is notably no suggestion that Congress lacked authority to delegate to the PTO the power to issue patents in the first instance. It would be odd indeed if Congress could not authorize the PTO to reconsider its own decisions.

The Board’s involvement is thus a quintessential situation in which the agency is adjudicating issues under federal law, “Congress [having] devised an ‘expert and inexpensive method for dealing with a class of questions of fact which are particularly suited to examination and determination by an administrative agency specially assigned to that task.’” *Stern*, 131 S.Ct. at 2615

This characterization badly misreads *Stern*, as the context from which the sentence fragment “[t]he patent right ‘derives from an extensive federal regulatory scheme’” is extracted makes clear. The first point to note is that the words “patent right” nowhere appear in that sentence or indeed anywhere in the opinion; nor, for that matter, does the word “extensive” appear anywhere in *Stern*. The exact words are “a federal regulatory scheme.” The full passage makes it clear that the Supreme Court sought to limit the scope of the public rights doctrine:

Shortly after *Northern Pipeline Construction Co. v. Marathon Pipe Line Co.*, the Court rejected the limitation of the public rights exception to actions involving the Government as a party. The Court has continued, however, to limit the exception to cases in which the claim at issue derives from a federal regulatory scheme, or in which resolution of the claim by an expert government agency is deemed essential to a limited regulatory objective within the agency’s authority.

The opinion goes on to say, citing *Bakelite*, “the [public rights exception] extended ‘only to matters that historically could have been determined exclusively by’ the Executive and Legislative
Branches,” which of course excludes all common law and equitable actions involved in patent litigation. The Court then cited its own earlier 2011 decision in United States v. Jicarilla Apache Nation, which relied on the initial public rights distinction in Murr v. Lessee, to support the application of the public rights doctrine applied to the trust obligations of the United States to the Indian Tribes because they “are established and governed by statute rather than the common law, and in fulfilling its statutory duties, the Government acts not as a private trustee but pursuant to its sovereign interest in the execution of federal law.”

There is not the slightest hint anywhere in Stern that removing all patent litigation from the federal courts counts as “a limited regulatory objective,” when it is in fact the most ambitious overhauling of the framework for patent adjudication ever. The government’s brief likewise elides all these difficulties, never once mentioning the dominant place that the Chief Justice accorded to Article III courts.

II. Due Process Concerns

The second issue that is raised by the IPR proceedings at stake in Oil States requires less discussion. Due process requires that parties have their cases resolved before neutral tribunals that are free from any form of bias or favoritism. Its purpose is to ensure that, in the words of Justice Thurgood Marshall in Marshall v. Jerrico, Inc., “no person will be deprived of his interests in the absence of a proceeding in which he may present his case with assurance that the arbiter is not predisposed to find against him.” There is not the slightest hint anywhere in Stern that removing all patent litigation from the federal courts counts as “a limited regulatory objective,” when it is in fact the most ambitious overhauling of the framework for patent adjudication ever. The government’s brief likewise elides all these difficulties, never once mentioning the dominant place that the Chief Justice accorded to Article III courts.

Recently, in PHH Corporation v. Consumer Financial Protection Bureau, now on rehearing en banc at the DC Circuit, Judge Brett Kavanaugh held that the principles of the separation of powers require that the President be able to remove the head of the Consumer Financial Protection Bureau, not only for cause, as the legislation provides, but also at will, in order to prevent the excessive concentration of power in the hands of a single unelected person who could design and institute his own regulations within his own agency. A similar concern arises when administrative bodies create administrative judgeships. The SEC’s internal tribunals were recently rebuffed under the Appointments Clause in Bandimere v. SEC because, the 10th Circuit held, SEC judges serve as “inferior officers” of the United States who, under the Appointments Clause in Article II, § 2, cl. 2 of the Constitution, may only be appointed by some “head of department”; because


In the patent context, procedural due process is denied when the head of the PTO may appoint to a particular case whomever he or she sees fit to appoint, with no external constraint. Judge Dyk, joined by Judge Wallach, wrote in a related case:

[W]e are also concerned about the PTO’s practice of expanding administrative panels to decide requests for rehearing in order to “secure and maintain uniformity of the Board’s decisions.” Here, after a three-member panel of administrative judges denied petitioner Broad Ocean’s request for joinder, Broad Ocean requested rehearing and requested that the rehearing be decided by an expanded panel. Subsequently, “[t]he Acting Chief Judge, acting on behalf of the Director,” expanded the panel from three to five members, and the reconstituted panel set aside the earlier decision.

Nidec [the objector] alleges that the two administrative judges added to the panel were chosen with some expectation that they would vote to set aside the earlier panel decision. The Director represents that the PTO “is not directing individual judges to decide cases in a certain way” (Director Br. 21). While we recognize the importance of achieving uniformity in PTO decisions, we question whether the practice of expanding panels where the PTO is dissatisfied with a panel’s earlier decision is the appropriate mechanism of achieving the desired uniformity. But, as with the joinder issue, we need not resolve this issue here.

The puzzling question is why the federal courts bother to tarry at all. The PTO process under the AIA reeks of opportunism and bias and begs to be struck down for denying due process.

III. Conclusion

The ambitious provisions found in the AIA upend the historical practice of patent litigation in ways that have already led to a call for major reforms. There are strong theoretical reasons to adhere to the traditional verities of litigation in these patent cases. The public rights doctrine has never been applied, in England or the United States, to disputes at common law or in equity where damages, an injunction, or both are sought. The rights in question in patent disputes are not created by the government through its examination and registration system. That system is simply intended to protect good patents, which arise from the labor, energy, and intelligence of the patentee. It is not intended to let the government act as if its certification of a patent were equivalent to creation of the patent right. To accept the underlying theory of the AIA is to make any property that is protected by a system of registration—which is virtually all real and intellectual property—a federal resource rather than a private resource. This would be a huge nationalization of private rights, which would inevitably lead to the pathologies that always impact government ownership of resources. The position that patent rights are private rights and that disputes must therefore be adjudicated in Article
III courts has been uniformly accepted in the United States from Murray's Lessee to Matal v. Tam. The need for Article III courts to resolve ordinary disputes between private parties is not just an accidental feature of the legal system. It represents an important constitutional safeguard, rooted in the separation of powers, that should not be frittered away by clever artificial distinctions and subtle arguments. The ordinary processes of adjudication have resolved patent disputes since the beginning of the Republic. There is no reason to replace them with an untested set of administrative procedures that can easily lead to partisanship and error that will undermine national confidence, not only in the patent system, but in the rule of law.
The Supreme Court Tackles Patent Reform: A Series of Articles Examining Oil States Energy Services, LLC v. Greene’s Energy Group, LLC

By Richard A. Epstein

Note from the Editor:
This article provides alternative answers to some of the questions posed at oral argument in Oil States. It is the second in a series of articles by Professor Epstein about the issues at stake in Oil States Energy Services, LLC v. Greene’s Energy Group, LLC. Earlier versions of these articles were posted on the Fed Soc Blog.

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.


About the Author:
Richard A. Epstein is the Laurence A. Tisch Professor of Law at NYU School of Law, the Peter and Kirsten Bedford Senior Fellow at the Hoover Institution, and the James Parker Hall Distinguished Service Professor of Law Emeritus and Senior Lecturer at the University of Chicago.

My thanks to Bijan Aboutorabi, Tom Molloy, and Max Semels, University of Chicago Law School students, for their excellent research assistance.

Post-Argument Article: Further Reflections on the Oil States Case after Oral Argument Before the Supreme Court

On November 27, 2017, the United States Supreme Court heard oral argument in Oil States Energy Services, LLC v. Greene’s Energy Group, LLC on the thorny question of whether the inter partes review (IPR) provisions of the America Invents Act (AIA) run afoul of constitutional guarantees on three related issues: separation of powers, the doctrine of unconstitutional conditions (i.e. whether Congress can condition the grant of a patent on an applicant agreeing to waive his or her Seventh Amendment right to a jury trial), and the due process clause of the Fifth Amendment.

The conventional wisdom is that the Court is likely to affirm the decision below (perhaps with dissents from Justices Roberts, Alito, and Gorsuch) and allow the Patent and Trademark Appeal Board (PTAB) to continue in its current form. Nevertheless, the oral arguments in the case expose deep and pervasive misconceptions as to both the desirability and the constitutionality of the AIA and IPR.

It is important to examine four related questions as we consider the oral arguments and what they suggest about the constitutional issues at stake in the case. First, when, if ever, does the PTAB lose its ability to claim jurisdiction over a case bought by a patent holder against an alleged infringer after the Patent and Trademark Office (PTO) grants the patent? Second, can the PTAB condition a grant of a patent on the patentee’s agreement that the PTAB retains control over future litigation over the validity of that patent, even in suits that the patent holder brings in federal district court? Third, is it consistent with due process for the Chief Judge of the PTAB to unilaterally add judges to any ongoing patent panel in order to promote uniformity in the issuance of patents—a power the PTAB has explicitly claimed? Fourth, if a potential infringer of a state-held patent asks for IPR, can the state be required to submit to the PTAB’s jurisdiction? Unfortunately, the questions from the Justices and the lawyers’ responses raise troublesome questions that do not admit of easy answers.

I. Patent Grants and the Timing of Inter Partes Review

The first challenge directed to Allyson Ho, counsel for Oil States, was aggressively stated by Justice Ruth Bader Ginsburg:
“there must be some means by which the Patent Office can correct the errors that it’s made, like missing prior art that would be preclusive.” (Tr. 3–4). Justice Kagan quickly chimed in by asking, as if the question had no real answer, where it is best to “draw the line” between those cases where traditional judicial procedures are required and those in which PTO procedures like IPR are both appropriate and efficient. (Tr. 17). Ho’s answer was that “trial-like” procedures cannot be upended through truncated procedures before the PTAB. (Tr. 19).

I would answer the questions posed by Justices Ginsburg and Kagan in a somewhat different manner. To Justice Kagan’s query on where to draw the line, the best answer is the same that it has always been: the patent office loses control over the patent once it is finally issued, so IPR-like procedures are only appropriate until the patent issues, after which all subsequent disputes must be resolved in court. That line is clear, sensible and known to all. Any line after the patent issues will be an exercise in arbitrary decision making. On this point, Justice Breyer sounded a sensible note of caution when he asked whether IPR could be triggered some ten years after a patent was granted when a company had spent some $40 billion in reliance on the grant. (Tr. 29). If the PTO always needs breathing room to correct its prior mistakes, the passage of time becomes a mere detail that warrants no consideration. This novel approach would reduce the venerable statute of limitations to an inconvenient obstacle that could be shunted aside in the endless pursuit of perfect justice.

In fact, the best answer to Justice Kagan’s question is that the correct line is that drawn in the pre-AIA precedents—at the time the grant of the patent becomes final. Benjamin Christoff summarizes the relevant cases as follows:

In McCormick Harvesting Machine Co. v. Aultman [169 U.S. 606 (1898)], the Supreme Court described in no uncertain terms the limitations on the PTO’s authority: “It has been settled by repeated decisions of this court that when a patent has [been granted] . . . it has passed beyond the control and jurisdiction of that office, and is not subject to revocation or canceled by the president, or any other officer of the government.” After issuance, the patent “has become the property of the patentee, and as such is entitled to the same legal protection as other property.” The Court later reiterated this view in Crown Cork & Seal Co. v. Ferdinand Gutmann Co. [304 U.S. 159 (1938)]: “After a patent is granted it passes ‘beyond the control and jurisdiction’ of the Patent Office; the proceedings are closed and the application cannot be amended nor serve as the basis for a new ‘divisional’ or ‘continuing’ application.”

Using the time of the completed conveyance to determine when any grant vests supplies a uniform theory for all government grants that is easily identifiable and enforceable. It also prevents the needless proliferation of different rules for different asset classes, and it prevents the systematic encroachment by Congress and the administrative state on the prerogatives of the judicial branch. Against these virtues of tradition and clarity, the PTO argued that the new procedures boast efficiency, costs savings, PTO expertise in adjudication, and gains to investor confidence.6 These purported justifications for allowing a PTAB review of issued patents are not persuasive.

On the first question of efficiency, the correct inquiry starts with a comparative analysis of how the PTO will work under the two alternative legal regimes. According to the conventional wisdom, the power to first revisit and then revise patents under the AIA is intended to “improve patent quality and restore confidence in the presumption of validity that comes with issued patents” because that review “screen[s] out bad patents while bolstering valid ones.”7 There is no empirical evidence on the strength of a novel procedure, so this claim rests on theory. But even economic theory offers no support for the proposition that the AIA procedure reduces both kinds of error—allowing bad patents and denying good ones.8 The key insight here is that the PTO is far more likely to take care in its initial examinations of patents if it understands that, once it grants a patent, its powers of review are over. This observation is not inconsistent with the view that some pre-grant procedure should allow challengers to come forward before any patent issues. The anticipated quality of the patent should be higher if the initial review is done expeditiously. The more thorough review at the first stage reduces the need for adjudication after that grant is made, which in turn lowers the level of system-wide uncertainty. The more expeditious process therefore cuts down on the cost of subsequent deliberations. The expertise of the PTO is respected, which is why the law creates a presumption in favor of patent validity.9 There is little reason to think that the expertise of the PTAB, as opposed to the entire PTO which initially awarded the patent, will be better a second time around, especially if the judging panel can be stacked to reach a desired result. Given the widespread distrust of the PTAB among investors, why expect that the IPR will restore or enhance investor confidence? Small inventors were the strongest opponents of IPR because they feared being bulldozed by large corporations who could now drag them and their assignees before the PTAB. Private investors want to invest in new technologies, not in multiple rounds of litigation.

Investor confidence will be decreased—not increased, as PTO supporters claim—by a distinctive feature of the AIA that allows the PTAB to continue with the IPR even after the private dispute between the patent holder and the alleged infringer.

6 See id.


9 35 U.S.C. § 282(a) (“A patent shall be presumed valid. . . .”).
has been settled. (Tr. 7). IPR thus places the patent holder in the difficult position of having two opponents instead of one. The PTAB readily strikes down patents on grounds of patent eligibility or prior art. The credible threat to institute an IPR before the hostile PTAB is likely to induce the patentee to settle the case on terms favorable to the challenger before it is brought back to the PTAB, which can still seek to invalidate the patent, thereby affording additional protection against suit to all other actual and potential infringers. This extraordinary feature is not found in ordinary trials, which supplies yet another reason the PTAB should not be given the unprecedented power that fuses administration and adjudication without the protections that attend real federal court disputes.

II. Unconstitutional Conditions

The ability to grant or deny a patent lies within the exclusive province of the PTO. Its legal monopoly in this area should trigger the application of the standard limitations that apply to any legal monopoly. It must be prevented from abusing its monopoly power so that all of the restrictions it imposes on patentees (like all other monopoly impositions) are consistent with the general objective of advancing overall social welfare. That conclusion follows from a proposition announced by Sir Matthew Hale, who said of wharves of advancing overall social welfare. That conclusion follows from a proposition announced by Sir Matthew Hale, who said of wharves


13 The same condition does not attach, however, to the behavior of individual patentees in the use of their patents. The patent is only an exclusive right to sell a given product, which never precludes new entrants from developing alternative devices to supply the same or superior functionality, so the concerns associated with monopoly power are not present. Illinois Tool Works Inc. Independent Ink, 547 U.S. 28 (2006) (holding that “the mere fact that a tying product is patented does not support such a presumption [of market power.”]).


15 Id. at 541. I do not think that this argument holds in the employment context in Loudermill, because labor markets are competitive and teaching positions are widely available from other employers. See Richard A. Epstein, Bargaining with the State at 245-246. But it does hold for patents given that the PTO and the PTAB both exercise monopoly power.
to a stronger political party, which is the essence of the political abuse associated with monopoly power.

The earliest cases involving the application of the doctrine of unconstitutional conditions were tied to a vital question of federal jurisdiction: the right of any individual or firm to avail itself of federal diversity jurisdiction, by bringing (as a plaintiff) a suit in federal court against a citizen of another state, or by removing (as a defendant) a case into federal court when sued by a defendant of a different state. The federal government can, of course, place some limitations on diversity jurisdiction, including minimum rights of suit or the exclusion of certain categories of cases. But it is a different kettle of fish to require a foreign corporation that seeks to do business in the state to forfeit all access to federal courts under diversity jurisdiction. Such a requirement was struck down in two early Supreme Court cases that are unquestioned today: *Insurance Co. v. Morse* and *Terral v. Burke Construction Co.*

These decisions were quite emphatic in denying the power of any given state to tie permission to do business in the state with the forfeiture of access to federal courts under diversity jurisdiction. In *Morse*, Justice Hunt wrote:

> Every citizen is entitled to resort to all the courts of the country, and to invoke the protection which all the laws or all those courts may afford him. A man may not barter away his life or his freedom, or his substantial rights. . . . In these aspects any citizen may no doubt waive the rights to which he may be entitled. He cannot, however, bind himself in advance by an agreement, which may be specifically enforced, thus to forfeit his rights at all times and on all occasions, whenever the case may be presented.19

And in *Terral*, Chief Justice Taft followed suit by writing:

> [A] State may not, in imposing conditions upon the privilege of a foreign corporation’s doing business in the State, exact from it a waiver of the exercise of its constitutional right to resort to the federal courts, or thereafter withdraw the privilege of doing business because of its exercise of such right, whether waived in advance or not.20

Any principled defense of freedom of contract is always subject to reasonable limitations when a state or a private party exercises monopoly power. In this instance, the Framers of the Constitution thought that access to diversity jurisdiction was sufficiently important to build its protection into the Constitution. The states cannot undermine that structural arrangement by conditioning the access of foreign corporations to the state on waiving that constitutionally protected option. If such a condition were allowed, most firms would waive the benefit of diversity jurisdiction, for the loss of access to federal courts in some future litigation is a small price to pay for the opportunity to do business in another state. But then every state would demand that concession and virtually every corporation would accede, all of which would amount to an implicit removal of the structural protection of diversity jurisdiction. Avoiding just this debacle is why the unconstitutional conditions doctrine retains its great tenacity.

Contrary to the remarks of the lawyers Kiss and Stewart, this doctrine also applies to the congressional power over patents. No inventor can afford to refuse to take a patent because its grant is conditioned on the waiver of various constitutional rights to adjudication, complete with jury trials, in an Article III court. The AIA is not exempt from this basic analysis of unconstitutional conditions. There is a fundamental right to have a case heard before a neutral tribunal. Under the Seventh Amendment, there is also a constitutional right to have questions of fact raised in suits at common law before a jury, in patent cases as in others. The only way for the government to resist the application of the unconstitutional conditions doctrine is to insist that there is no independent constitutional right to a jury trial in patent cases.

It is of course possible to deny that the right to a jury trial exists in patent cases, as Mark Lemley has recently done in his article *Why Do Juries Decide if Patents are Valid?*21 Lemley’s basic position is that, “while patent lawyers today take for granted the power of the jury to decide whether the PTO made a mistake in issuing a patent, the role of the jury in patent cases is a recent and unusual phenomenon with a murky history.”22 In his view, the decisions of administrative agencies normally are not subject to second-guessing by juries. But in the patent context, the decisions in question are in fact grants, and on these the history is far clearer. As H. Tomás Gómez-Arostegui and Sean Bottomley have shown in their amicus curiae brief in *Oil States*, Lemley’s historical account is riddled with errors, because in fact courts of equity routinely sent cases to the law courts for final disposition in which all questions of fact were resolved by juries. After an exhaustive study, they conclude that:

> The cases demonstrate that juries regularly decided the following issues related to validity:

1. whether the invention was new;
2. whether the patentee was the actual inventor of the purported invention;
3. whether the invention was useful;
4. whether the specification accurately described the claimed invention; and
5. whether the specification enabled a person working in the relevant art to construct the item described in it.23

The question of patent eligibility is not clearly included in that list, but it is just that determination that has to be made to hold

---

1. 87 U.S. 445 (1874).
6. Id. at 1674.
that something is an invention, as opposed to either an idea or a product of nature. The question of patent validity was not, as Lemley argues, a mix between legal and equitable considerations inconsistent with the general relationship between the two sides of the courts. In oral argument, Justice Sotomayor, citing the brief of Gómez-Arostegui and Bottomley, made much of the position of the English Privy Council, which had long operated chiefly as an advisory body to the Crown, but which often served as the appellate court of last resort for legal disputes that arose in the colonies or former colonies. In addressing this issue, Justice Sotomayor observed “the fact that [the influence of the Privy Council] waned didn’t mean that it was eliminated, and it didn’t mean that the Privy Council or the crown thought that it no longer had those rights.” (Tr. 25-26). But as Gómez-Arostegui and Bottomley make clear, the fading jurisdiction of the Privy Council never reached the issues set out above. They first state explicitly that the role of the Privy Council had been “inadvertently overstated” by many scholars who had not looked at the full range of available sources. More to the point, they noted: “As a forum to adjudicate infringement, the Council played essentially no role after the Restoration [i.e. 1660], with enforcement instead falling to the Chancery, King’s Bench, Common Pleas, and Exchequer.”

To be sure, the Privy Council retained a rarely used process that was much more cumbersome than the truncated AIA processes at issue in Oil States. In the Privy Council, no one person could decide the issue. The initial petition was first referred to a committee that in turn delegated the matter to the Attorney General, who used affidavits and testimony to examine the petition, after which it could be voided only with the concurrence of six other members of the Privy Council. The entire process included many veto points on the road to invalidation. It would be odd to treat this rare and complex process as the decisive precedent to legitimate the wholly different invalidation process that under the AIA is vested virtually in a single person. The accurate guide in this case is the Statute of Monopolies, which made it clear that “the force and validity of [patents], and every of them, ought to be, and shall be for ever hereafter examined, heard, tried, and determined, by and according to the common laws of this realm, and not otherwise.”

III. DEPRIVATIONS OF DUE PROCESS

The third major concern raised in Oil States is whether the right of the Chief Judge of the PTAB to add new members to a panel during the course of a hearing comports with the notions of due process that demand a hearing before a neutral and informed judge. The issue of panel assignments is important in all appellate courts, and a recent article by Professor Marin K. Levy offered an exhaustive review of these practices in the federal courts of appeals. The premise of her study is that most circuit courts announce that they make random assignments of judges to individual cases. The obvious justification for this practice is that it prevents any effort to stack a panel with judges who are inclined to a given point of view. Levy summarizes her position as follows: “Although the courts generally tried to ‘mix up’ the judges, the chief judges and clerks responsible for setting the calendar also took into account various other factors, from collegiality, to efficiency-based considerations.” At no point does Levy’s study make any reference to a process whereby the chief judge of a circuit can stack the tribunal with persons whose views are sympathetic to his or her own. The notion of due process—which traditionally involves a trial before a neutral panel under fixed and determinate rules—is mocked when the PTAB is allowed to stack a panel with sympathetic judges, contrary to the practice of every other court.

Before the Supreme Court, both Kise and Stewart downplayed the importance of this extraordinary power. Mr. Kise said that “I don’t believe that’s taken place more than one or two times.” (Tr. 33-34). He was cut off before he could say that he did not think it took place here, but then in response to Justice Kennedy’s question said, “if it were rampant, then I think what the Court said in Cuzzo...” that the Administrative Procedure Act and other provisions of the Constitution would deal with infirmities in a particular case on an as-applied basis.” (Tr. 33-34). The matter was then left hanging until, in response to a question by Chief Justice Roberts—“Does it comport with due process to change the composition of an adjudicatory body halfway through the proceeding? (Tr. 45)—Mr. Stewart offered this response: “This has been done on three occasions. It has been done at the institution stage,” (Tr. 45), i.e. at the outset of the case and not in the middle. This exchange understates the seriousness of the problem. If the issue were indeed one that should be resolved on an “as-applied basis,” the due process challenge should succeed in all such cases. Hence these are really facial challenges which are more efficiently addressed at the wholesale level, unless some argument could be put forward to explain that the practice makes sense in some cases, but not in others. But there is no hint anywhere as to what the relevant criteria for carving out this exception should be. Nor, as a matter of fact, does it appear that these events are as infrequent as was claimed, for there were at the time of oral argument many pointed charges of improper conduct.

In a deeper sense, however, the frequency of use should not matter, because it turns out that David Ruschke, Chief Judge of the PTAB, has made panel stacking official policy:

The PTAB’s chief judge noted that there are four instances in which the agency’s standard procedures give him the

---

25 Id. at 34.

26 Statute of Monopolies, 1624, 21 Jac. I, c. 3.

27 For my earlier comments on this point, see Epstein, Tackles, supra note 8, at sec. II.


29 Id. at 65.

30 136 S. Ct. 2131, discussed infra at Section I.

31 See Pfeifer, supra note 4 at 1. See also Bill Smith, Federal Circuit’s Concern Regarding PTAB ‘Panel Stacking’—Back to the Future?, IP INTELLIGENCE, August 29, 2017 (citing Nidec Motor Corp. v.
authority to expand a panel on a validity trial: when the trial involves an issue of exceptional importance; to maintain uniformity of PTAB decisions; when the USPTO Director submits a written request on an issue of first impression; or when the director feels that the panel should not follow a previous board decision to better serve the public interest.

However, Ruschke noted that his authority to expand the panels for PTAB trials doesn’t require him to notify the parties in the trial that the decision to expand the panel has been made. . . .

“I’ve only expanded [the panel] in situations where I’ve added [Vice Chief Judge] Scott [Boalick] and myself to emphasize a unanimous decision below,” Ruschke stated.32

The astounding assertion of power was made on November 16, 2017, eleven days before the oral argument in Oil States, in an executive session held to plan future PTAB policy. Ruschke publicly stated that position two days after the oral argument in which Kise and Stewart made apparently inconsistent representations. This is no longer a matter of speculation, for on December 19, 2017, the PTAB in Ericsson Inc. v. University of Minnesota33 announced in a judgment in which both Ruschke and Boalik sat that “our standard operating procedures provide the Chief Judge with discretion to include more than three judges,” claiming that this power was delegated expressly to the Chief Judge by the AIA under In re Alappat.34

But the precise question in Alappat was this: “When a three-member panel of the Board has rendered its decision, does the Commissioner have the authority to constitute a new panel for purposes of reconsideration?”35 The case does not mention any power to reconsider if the initial panel has not rendered its final judgment on the issue. And the difference is palpable, for under the procedure sanctioned in Alappat, any reviewing court gets to read two opinions which come to different conclusions. Nor does Ericsson comment on the possible reasons for allowing a second panel to reconsider a case. The point is uncertain, but Alappat reads like a provision calling for an en banc hearing that allows for a broader range of views, not for the particular programmatic agenda of the Chief Judge. And even if Alappat is read that broadly, its decision is not binding on the Supreme Court. At the very least, Ericsson only magnifies the uneasiness about PTAB procedures, because its decision was announced after oral argument in Oil States. The Supreme Court should be apprised of this development by all or some of the parties before it is asked to render its final decision.

PTAB’s public statement of its panel-stacking procedure flouts due process requirements. The four stated reasons for expanding panels are themselves good reasons not to allow the Chief Judge to tamper with panel composition, let alone make himself the decisive voice on the case. Irregularities of procedure are least welcome in cases of “exceptional importance.” It is precisely in these cases that no one person should be allowed to wield inordinate influence. More specifically, unlike en banc hearings, which are intended to collect a wide range of disparate views, this procedure involves the Chief Judge appointing only himself and the Vice Chief Judge to panels for the express purpose of rigging the outcome in favor of the legal positions that these two judges favor. Doing this in secret only adds insult to injury. But whether done openly or in secret, placing such inordinate power in the hands of one or two individuals is an open admission of prejudice that should never be allowed in any tribunal, whether an Article III court, which has clear functions, or the PTAB which is an unruly amalgam of judicial and administrative functions. It is the sign of a broken culture that public officials should trumpet their own misconduct as a form of public virtue.

IV. SUBSTANTIVE ISSUES IN THE PTAB

The PTAB’s control extends to a full range of procedural and substantive issues. It is far from clear that the PTAB should have dispositive power in many of these areas, given that its decisions tend always to strengthen its dominant ethos that strong patent rights are a mistake. Michelle Lee was the Director of the PTO when Cuozzo Speed Technologies, LLC v. Lee was decided;36 she joined the agency after her tour of duty with Google, which is known for its general anti-patent positions. Up for decision in that case were two issues: appeal and claim construction.

The initial question was whether the PTO Director’s decision “whether to institute an inter partes review under this section shall be final and non-appealable”;37 the Supreme Court answered this question in the affirmative, given the clear language of the statute. The second question was whether the authority delegated to the PTO under the AIA to issue regulations “governing inter partes” is broad enough to uphold a PTO regulation that provided that in the course of IPR any patent claim “shall be given its broadest reasonable construction in light of the specification of the patent in which it appears.”38 The standard judicial rule calls for the “ordinary meaning” test to apply.

A. Appealability

On the first of these rulings, it is generally the case that parties do not get an appeal from any decision to institute an action. But in this case, that power precludes the right of the parties to end the case by a mutual settlement of their differences, so it is at least an open question whether the patent holder is entitled to have some escape from a process that he

---

33 Ericsson v. Regents of the University of Minnesota, IPR2017-01186, -01197, -01200, -01213, -01214, and -01219.
34 33 F.3d 1526 (Fed. Cir. 1994).
35 Id. at 1530.
36 136 S. Ct. 2131.
38 37 C.F.R. § 42.100(b).
cannot terminate by ordinary means. In Cuozzo, the Court held that the statutory language was not limited to a denial of interlocutory appeals, but covered decisions to initiate cases as well. Clearly, allowing appeals on an interlocutory basis is rarely wise, but the ability to stop a misguided suit before it begins has much more to commend it because it does not involve breaking up an ongoing suit. Nonetheless Justice Breyer, over dissent, held that the clear language of the statute covered the routine case before it. Accordingly, he rightly set aside cases that present constitutional issues or whose questions went outside the scope of the particular section at issue. The decision is debatable, but defensible. But the result should be exactly the opposite whenever the Chief Judge exercises his power under the statute to stack the panel in favor of his preferred outcome, for there is no reason a party should have to try a case before a biased panel to the bitter end when the objection is to the method by which the panel is constituted. In these cases, the ideal solution is a per se ban against this practice. But if the matter becomes an as-applied challenge, it should be resolved before anyone has to go through a proceeding that could well be invalidated because of problems at the outset of the proceeding.

B. Claim Construction

Justice Breyer’s decision in Cuozzo is, however, far more troublesome for its rote invocation of Chevron, which in this instance allows the PTAB to fashion rules of claim construction, displacing the traditional rule of ordinary construction by an untested rule that allows the PTAB to give claims the broadest reasonable interpretation. At this point, Justice Breyer again observes that the PTO may intervene in appellate decisions, even if the private challengers do not join in the appeal, without quite saying that the appellate court would be bound to follow Chevron, and thus give deference to the PTAB in cases of statutory ambiguity. Under the AIA, the burden of proof on the challenger, the PTO, or both is a bare preponderance of evidence instead of the higher standard of clear and convincing evidence, which only increases the incentive for patent challengers to avoid the district courts. And the broader claim reading makes it more likely that any patent will be invalidated because it is more likely to infringe on a preexisting patent, for if it would not the parties would normally have opted for the broader construction anyhow.

Any set of rules that abandons ordinary construction should be viewed with deep suspicion. Yet the justification for using one standard before the PTAB and another in the federal courts is utterly unconvincing. Justice Breyer insists that, “in addition to helping resolve concrete patent-related disputes among parties, inter partes review helps protect the public’s ‘paramount interest in seeing that patent monopolies . . . are kept within their legitimate scope.’” The ellipsis in the quotation conceals the fact that the case cited, Precision Co. v. Automotive Company, was concerned with the state's “paramount interest in seeing that patent monopolies spring from backgrounds free from fraud or other inequitable conduct, and that such monopolies are kept within their legitimate scope.” No issue of claim construction or monopoly power was raised in Precision. It is therefore critical to ask why a novel test of claim construction is superior to the traditional test when the state also has a powerful interest in preserving incentives to invent by ensuring that inventors receive a just return on their inventions. Put otherwise, the social welfare and private incentives are well aligned under the traditional rule. Yet the Breyer opinion offers no word of explanation as to why a major deviation from that standard should be accepted. Cuozzo thus gives the PTAB the power to redraw substantive law, without any explanation of whether or why a reviewing court should accept its determination on that point. There is no such embarrassment if the disputed AIA provision is read to allow the PTAB to develop its own procedures to administer the standard legal rule. This decision converts the PTAB into a runaway train.

C. Sovereign Immunity

One of the most litigated questions in recent years concerns the ability of the state to invoke the defense of sovereign immunity against tort actions brought by private parties. That issue was not raised in Oil States, but the PTAB’s decision in Ericsson—which involved a state university as patentee—was handed down two days after oral argument in Oil States. The aggressive position taken in Ericsson—that the state cannot invoke the defense of sovereign immunity against a motion of the patent challenger to return the case to the PTAB—once again illustrates the inordinate power that the PTAB is arrogating to itself. To put the issue in context, the current law affords states immunity against suits by private parties, but not by state governments. It also holds that a state may voluntarily waive its defense of sovereign immunity under the Eleventh Amendment to allow suits to go forward. The more difficult branch of sovereign immunity arises in those cases

---

39 Cuozzo, 136 S. Ct. at 2140. The dissent of Justice Alito joined by Justice Sotomayor is found at 2149.
40 Id. at 2136, 2142-2144.
41 35 U.S.C. § 316(c).
42 For the dangers in of abandoning ordinary construction in connection with defamation, see the application of the innocent construction rule in Loos v. Levits, 556 F.3d 564, 567 (7th Cir. 2009) (dismissing a defamation suit for false charges of “fabrication” as “a description of an academic dispute regarding controversial theories, not an accusation of academic dishonesty”).
43 Cuozzo, 136 S. Ct. at 2144.
44 324 U.S. 806, 816 (1945).
45 Sovereign immunity extends to suits against a state brought by the state's own citizens. See Hans v. Louisiana, 134 U.S. 1 (1890). That decision is technically not under the Eleventh Amendment, but rather relies on the general structural considerations set out by Alexander Hamilton in Federalist 81:

It is inherent in the nature of sovereignty not to be amenable to the suit of an individual WITHOUT ITS CONSENT. This is the general sense, and the general practice of mankind; and the exemption, as one of the attributes of sovereignty, is now enjoyed by the government of every State in the Union. Unless, therefore, there is a surrender of this immunity in the plan of the convention, it will remain with the States.

47 U.S. Const., amend. XI (“The judicial power of the United States shall not be construed to extend to any suit in law or equity, commenced or
in which a state is deemed to have waived its sovereign immunity because of actions that it has taken prior to litigation. In *Lapides v. Board of Regents of University System of Georgia*, the plaintiff sued the University of Georgia, alleging that it had improperly placed allegations that he had engaged in sexual harassment in his personnel file, in violation of state and federal law. The state removed the case to federal court, and the Supreme Court held that it thereby waived the defense of statutory immunity on the ground that it could generate unfair results to allow the state to first avail itself of federal jurisdiction only to then plead sovereign immunity. It further noted that:

[A]n interpretation of the Eleventh Amendment that finds waiver in the litigation context rests upon the Amendment’s presumed recognition of the judicial need to avoid inconsistency, anomaly, and unfairness, and not upon a State’s actual preference or desire, which might, after all, favor selective use of immunity to achieve litigation advantages.

That result does not apply in any patent dispute because those must be litigated in federal courts. The question, then, is why a state should be deemed to waive its immunity to IPR when it has no choice in where to go, given that the sole rationale in *Lapides* “was to prevent a State from selectively using its immunity to achieve a litigation advantage.” In order to patch the hole in its argument, the PTAB cites *New Mexico v. Knight*, which held that the states necessarily waived sovereign immunity against compulsory counterclaims, which had to be brought in the initial case or forever lost. But that precedent is completely inapposite in this instance. A state waives sovereign immunity to any compulsory counterclaim when it initiates a suit in federal court, but it hardly should be taken to waive the defense of sovereign immunity against IPR when it is not to its advantage to do so, and there is no reason for that to happen given that the federal district court offers a federal forum in which both sides compete on rough parity. Why would any state that has granted its opponent a federal forum ever throw itself into the lion’s den of the PTAB? There is accordingly no issue of fairness or strategic behavior that justifies the PTAB’s decision. Sadly, at no point did the PTAB opinion seek to reconcile its own opinion on Minnesota’s invocation of sovereign immunity with those advanced in the cases it cited or explain how the doctrine of sovereign immunity could ever come into play in cases where IPR is sought. Any case in which a tribunal such as the PTAB decides a case in favor of its own jurisdiction should be looked upon with deep suspicion, especially when the tribunal has never ruled in ways that have limited its power.

---

49  id. at 613.
50  id. at 620.
51 Ericsson, supra note 33, at 7.
52 312 F.3d 1111 (Fed. Cir. 2003).

V. Conclusion

The case law under the AIA is at a major turning point. The general perception in 2011 was that the AIA was making modest adjustments on a large number of technical points, with perhaps more profound changes on such matters as business method patents. It would be hard to find anyone who thought that the creation of IPR review would lead to a massive revolution in the operation of the entire patent system. But so it might be if the PTAB can control the agenda from top to bottom and transform the patent system into one that has the following features: stacked panels in favor of the views of the Chief Judge of the PTAB, no right to jury trial, a lower burden of proof to set aside a patent, no ability to settle out private disputes once the case is before PTAB, nonappealability of its decision whether to hear a case, a rule of claim construction that deviates from the well-settled rules of ordinary meaning, and an automatic waiver of sovereign immunity whenever a defendant removes a case from federal district court to the PTAB. These massive shifts are taking place without any showing that any of these reforms will do anything to improve the overall operation of the patent system. The traditional virtues of a system of separation of powers are made crystal clear by this litany of mistakes: the traditional regime lets the PTO decide which patents to issue, after which their protection is a judicial function. *Oil States* gives the Supreme Court the chance to stop a process that has already run off the rails. And if it does not, Congress should take steps to restore the proper constitutional balance.
The Supreme Court Tackles Patent Reform: A Series of Articles Examining Oil States Energy Services, LLC v. Greene’s Energy Group, LLC

By Richard A. Epstein

Note from the Editor:
This article criticizes Justice Clarence Thomas’ opinion in Oil States. It is the third in a series of articles by Professor Epstein about the issues at stake in Oil States Energy Services, LLC v. Greene’s Energy Group, LLC.

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.


About the Author:
Richard A. Epstein is the Laurence A. Tisch Professor of Law at NYU School of Law, the Peter and Kirsten Bedford Senior Fellow at the Hoover Institution, and the James Parker Hall Distinguished Service Professor of Law Emeritus and Senior Lecturer at the University of Chicago.

My thanks to Adam Mossoff, Professor of Law at Antonin Scalia Law School, George Mason University, for comments on an earlier draft, and to Nathaniel J. Tisa and Alexander Rubel, NYU Law classes of 2019 and 2020 respectively, for their excellent research assistance.

3 Oil States, 138 S. Ct. at 1380-86 (Gorsuch, J., dissenting).
4 This is a combination to which I have, it is no secret, been long opposed. See Richard A. Epstein & Kayvan B. Noroozi, Why Incentives for “Patent Holdout” Threaten to Dismantle FRAND, and Why It Matters, 32 BERKELEY TECH. L.J. 1381 (2017); Richard A. Epstein, The Disintegration of Intellectual Property: A Classical Liberal Response to Premature Obituary, 62 STAN. L. REV. 455 (2010).
9 See 37 C.F.R. §42.108(c) (2017).
effectiveness of the patent system, a process that accelerated with the ill-advised revisions of the Patent Act of 1952 by the AIA.\textsuperscript{10}

At issue in \textit{Oil States} were provisions of the AIA that authorize any party sued for infringement in federal court to initiate IPR before the PTAB after a final patent has been issued.\textsuperscript{11} Justice Thomas stresses the "narrowness of our holding," given that it only deals with IPR.\textsuperscript{12} But his flawed reasoning has powerful implications for the future direction of both the patent system and administrative law in general. As described by Professor Gregory Dolin, the AIA's procedures are the final stage of a transformation in patent practice that began more than three decades ago.\textsuperscript{13} Until 1981, administrative review was only allowed before a final patent had been issued. Once the patent issued, it counted as a completed grant that, like other grants of property, could only be modified or set aside in an Article III judicial proceeding, in which the challenger could show that the grant was improperly made or that the grantee had failed to comply with some of its terms. Historically, these standard rules of property applied with equal force to patents and land grants, since both are constitutionally protected forms of property.\textsuperscript{14} In 1981, as part of the Bayh-Dole Act, Congress adopted a procedure that allowed any member of the public to demand that the USPTO take a "second look" at the validity of any patent already issued.\textsuperscript{15} That review was not mandatory, and it was restricted to a reexamination on two key questions: novelty and nonobviousness. Nevertheless, it changed the fundamental balance between the Executive and the Article III courts, paving the way for the AIA's novel IPR practice. The 1981 reforms did not lead to a massive level of patent innovation, so Congress passed a second reexamination statute in 1999\textsuperscript{16} that persisted until its repeal and replacement by the AIA, which instituted the current form of IPR. \textit{Oil States} continues this trend by expanding the power of administrative agencies at the expense of Article III courts in adjudicating patent disputes, which has the effect of narrowing the intellectual property rights of patentees. I have already commented on how this decline in the doctrine of separation of powers will work to undermine the regime of stable property rights.\textsuperscript{17}

This article evaluates Justice Thomas' reasoning in \textit{Oil States} by following the structure of his opinion. In Part I, it examines and rejects his claim that patents fall "squarely" within the domain of public rights for which adjudication before an Article III court is not necessary.\textsuperscript{18} In Part II, it examines the effect of IPR on the patent system, with an emphasis on the separation of powers issues that \textit{Oil States} raises. It concludes with a brief comment on how various parties—including Congress and intellectual property practitioners—might address the problems with the patent system that the Court failed to correct in \textit{Oil States}.

I. Are Patents Public Rights?

A. \textit{Stern} and Northern Pipeline: Limiting the Domain of the Bankruptcy Courts

Justice Thomas begins his analysis with a discussion of the distinction between private rights, which are causes of action between parties, and public rights, which involve public issues in which the government has a direct stake. He starts by referencing the two key modern cases—\textit{Stern v. Marshall} and \textit{Northern Pipeline Construction Co. v. Marathon Pipe Line Co.}—to make two points: that the Court's public rights jurisprudence has not been "definitively explained," and that its precedents have "not been entirely consistent."\textsuperscript{19} But he does not mention that both cases emphatically denied the power of Congress to remove certain causes of action from the jurisdiction of Article III courts by declaring the rights at issue to be public rights. In \textit{Stern}, Chief Justice John Roberts stressed that the bankruptcy court—a non-Article III court—did not have "constitutional authority" to enter judgment on common law claims, including damages actions for patent infringement, which have been regarded as actions at common law since the Statute of Monopolies of 1623.\textsuperscript{20} He further added that "Congress may not 'withdraw from judicial cognizance any matter which, from its nature, is the subject of a suit at the common law...’."\textsuperscript{21} Under \textit{Stern}, suits at common law, including disputes over patent rights, must be adjudicated in Article III courts.

Earlier, in \textit{Northern Pipeline}, Justice William Brennan anticipated the centrality of Article III courts when he refused to allow their ouster by revised bankruptcy courts. After his own exhaustive account of the separation of powers, he concluded:

In sum, our Constitution unambiguously enunciates a fundamental principle—that the "judicial Power of the United States" must be reposed in an independent Judiciary. It commands that the independence of the Judiciary

\begin{itemize}
  \item \textsuperscript{10} Leahy-Smith America Invents Act, 35 U.S.C. § 66 et seq. (2012). For a critique of the AIA, see OS I, supra note 2, at 116-17.
  \item \textsuperscript{11} OS I, supra note 2, at 116-18. The USPTO defines inter partes review as "a trial proceeding conducted at the [PTAB] to review the patentability of one or more claims in a patent only on a ground that could be raised under §§ 102 [dealing with novelty and prior art] or 103 [dealing with nonobvious subject matter], and only on the basis of prior art consisting of patents or printed publications. United States Patent and Trademark Office. Inter Parties Review, Patent Trial and Appeal Board (May 9, 2017), https://www.uspto.gov/patents-application-process/appealing-patent-decisions/trials/inter-parties-review. The relevant passages from that definition are neither quoted nor analyzed.
  \item \textsuperscript{12} Oil States, 138 S. Ct. at 1379.
  \item \textsuperscript{13} Gregory Dolin, Dubious Patent Reform, 56 B.C. L. Rev. 881, 883-85, 896-97 (2015).
  \item \textsuperscript{14} See infra at sections I.E. & I.F.
  \item \textsuperscript{17} See OS I, supra note 2, at 117-18.
  \item \textsuperscript{18} Oil States, 138 S. Ct. at 1373.
  \item \textsuperscript{19} Id. at 1372 (citing \textit{Stern v. Marshall}, 564 U.S. 462, 488 (2011) and \textit{Northern Pipeline Constr. Co. v. Marathon Pipe Line Co.}, 458 U.S. 50, 69 (1982)).
  \item \textsuperscript{20} Statute of Monopolies, 1624, 21 Jac. I. c. 3, discussed in OS I, supra note 2, at 118.
  \item \textsuperscript{21} Stern, 564 U.S. at 484 (citing Murray's Lessee v. Hoboken Land & Improvement Co., 59 U.S. 272 (1856)).
\end{itemize}
be jealously guarded, and it provides clear institutional protections for that independence.\textsuperscript{22} He noted that “as an inseparable element of the constitutional system of checks and balances, and as a guarantee of judicial impartiality, Art. III both defines the power and protects the independence of the Judicial Branch.”\textsuperscript{23} That impartiality is wholly lacking in cases tried before PTAB judges; the Director of the USPTO can expand or alter the composition of a panel at will in order to reach a decision to his or her liking.\textsuperscript{24} By contrast, judges who sit on Article I courts dealing with taxation or bankruptcy have long, fixed terms and use neutral assignment systems to avoid the possibility of strategic misbehavior.\textsuperscript{25}

\textbf{B. Crowell and Murray’s Lessee: Distinguishing Public and Private Rights}

Next, Justice Thomas discusses a snippet from \textit{Crowell v. Benson}.\textsuperscript{26} \textit{Crowell} was a difficult case in which Chief Justice Charles Evans Hughes upheld the Longshoremen’s and Harbor Workers’ Compensation Act of 1927 against a constitutional challenge.\textsuperscript{27} This federal statute was modeled on state workers’ compensation systems that presented few difficulties under state constitutions. The hard question in \textit{Crowell} was whether their novel form met the more stringent federal due process and separation of powers requirements for Article III courts. The Supreme Court held that the employer must be allowed to obtain de novo review in district court on the factual jurisdictional question of whether the claimant was an employee of the defendant.\textsuperscript{28} The case was a delicate accommodation because it stretched the principles governing the constitutional position of Article III courts to fit a new structure—a more rapid and certain administrative workers’ compensation scheme that displaced factual resolution of tort claims in court—without abandoning judicial oversight altogether.

Justice Thomas writes, referring to \textit{Crowell}, “Our precedents have recognized that the [public-rights] doctrine covers matters which arise between the Government and persons subject to its authority in connection with the performance of the constitutional functions of the executive or legislative departments.”\textsuperscript{29} This sentence distorts Hughes’ opinion. First, it omits the beginning of the quoted sentence, inverting its meaning. The full sentence reads: “As to determinations of fact, the distinction is at once apparent between cases of private right and those which arise between the Government and persons subject to its authority in connection with the performance of the constitutional functions of the executive or legislative departments.”\textsuperscript{30} \textit{Crowell} came out as it did because Article III courts retained control over all questions of law. Chief Justice Hughes worried about how workers’ compensation claims were tried in the first instance before a commission on which no Article III judge sat. But he thought the statute overcame this difficulty by making it clear that “[r]ulings of the deputy commissioner [who oversees these cases] upon questions of law are without finality. So far as the latter are concerned, full opportunity is afforded for their determination by the federal courts through proceedings to suspend or to set aside a compensation order. . . .”\textsuperscript{31} \textit{Crowell}’s reasoning offers no warrant for the PTAB, which has virtually final authority over all questions of both fact and law insofar as it receives full deference from reviewing courts.\textsuperscript{32} There is no doubt that PTAB judges count as “inferior officers” under the appointments clause, which gives some indication of the breadth of their powers.\textsuperscript{33}

\textit{Crowell} points to Murray’s Lessee v. Hoboken Land & Improvement Co. to explain why the deputy commissioners—non-Article III judges—do not adjudicate public rights, further weakening Justice Thomas’ reliance on the case.\textsuperscript{34} Chief Justice Hughes first noted that a large number of cases dealing with “interstate and foreign commerce, taxation, immigration, the public lands, public health, the facilities of the post office, pensions and payments to veterans” involved public rights.\textsuperscript{35} In stark contrast, he concluded that the statutory scheme at issue in \textit{Crowell} could not be properly considered a public rights case. Accordingly, he held that workers’ compensation claims that concerned only ordinary individuals involved only private, not

\begin{itemize}
\item\textsuperscript{22} Northern Pipeline, 458 U.S. at 60.
\item\textsuperscript{23} Id. at 58.
\item\textsuperscript{24} See 35 U.S.C. § 6(c) (“Each . . . inter partes review shall be heard by at least three members of the Board, who shall be designated by the Director.”).
\item\textsuperscript{25} The U.S. Tax Court consists of 19 members appointed by the president for 15-year terms. 26 U.S.C. § 7443 (2018). Assignment of judges follows from their geographic assignment by the chief judge for each trial session—each judge is responsible for a portion of the Court’s national jurisdiction. Nicholas R. Merril & Mary W. Prosser, \textit{Litigating a Case Before the U.S. Tax Court}, The Federal Lawyer, 34-35 (Aug. 2014). Special trial judges may be assigned by the chief judge of the Tax Court to preside over small tax matters subject to special constraints and the ability of parties to seek review. 26 U.S.C. § 7443A; U.S. Tax Ct. R. 180-183 (July 2012).
\item\textsuperscript{26} 285 U.S. 22 (1932).
\item\textsuperscript{28} Crowell, 285 U.S. at 65.
\item\textsuperscript{29} Oil States, 138 S. Ct. at 1373 (citing Crowell, 285 U.S. at 50).
\item\textsuperscript{30} Crowell, 285 U.S. at 50 (emphasis added).
\item\textsuperscript{31} Id. at 45-46.
\item\textsuperscript{32} See OS II, supra note 2, 129-30 (citing Cuozzo, 136 S. Ct. at 2131). In several recent decisions, the Federal Circuit has given immediate issue-preclusive effect to its affirmation of PTAB findings of unpatentability. See, e.g., XY, LLC v. Trans Ova Genetics, L.C., 2018 U.S. App. LEXIS 13530, 2018 WL 2324460 (May 23, 2018).
\item\textsuperscript{33} See, e.g., Freytag v. Commissioner, 501 U.S. 868 (1991) (special trial judges of the Tax Court are inferior officers even though they do not have final authority); Lucia v. SEC, No. 17-130 (U.S. June 21, 2018) (same for SEC administrative judges).
\item\textsuperscript{34} Crowell, 285 U.S. at 50 (citing Murray’s Lessee, 59 U.S. at 284).
\item\textsuperscript{35} Id. at 51.
\end{itemize}
public rights. When these actions were brought at common law, Hughes continued, the right to jury trial was preserved in an Article III case by the Seventh Amendment, but in cases that arose either in equity or admiralty, Congress could change how factual issues can be resolved, so that the use of special masters and commissions was wholly proper.\(^\text{40}\) No one disputes this proposition; the only question is whether post-grant disputes over patent validity in an Article III court could be resolved, so that the use of special masters and commissions was wholly proper.\(^\text{16}\)

Crowell thus explicitly rejects the notion that Congress may conductive control over new statutory causes of action—whether they involve legal or equitable remedies—in an administrative tribunal instead of an Article III court. Instead, Article III gives Congress discretion only as to how factual issues should be resolved in equity and admiralty cases.\(^\text{37}\)

That result is borne out by recalling that Murray’s Lessee goes to great lengths to explain that Congress cannot encroach on the jurisdiction of Article III courts.\(^\text{38}\) Murray’s Lessee upheld an 1820 statute that provided for administrative resolution of a dispute over money that customs collectors owed to the United States; it did so because the Court found that the monetary dispute in question was in the nature of a public right. But Murray’s Lessee is easily distinguishable from Oil States. Murray’s Lessee sustained the statute against a due process challenge in part because the dispositive control over new statutory causes of action—whether they involve legal or equitable remedies—in an administrative tribunal instead of an Article III court. Instead, Article III gives Congress discretion only as to how factual issues should be resolved in equity and admiralty cases.\(^\text{37}\)

That result is borne out by recalling that Murray’s Lessee goes to great lengths to explain that Congress cannot encroach on the jurisdiction of Article III courts.\(^\text{38}\) Murray’s Lessee upheld an 1820 statute that provided for administrative resolution of a dispute over money that customs collectors owed to the United States; it did so because the Court found that the monetary dispute in question was in the nature of a public right. But Murray’s Lessee is easily distinguishable from Oil States. Murray’s Lessee sustained the statute against a due process challenge in part because the dispositive control over new statutory causes of action—whether they involve legal or equitable remedies—in an administrative tribunal instead of an Article III court. Instead, Article III gives Congress discretion only as to how factual issues should be resolved in equity and admiralty cases.\(^\text{37}\)

D. Franchise, Properly Understood

After misconstruing Bakelite, Justice Thomas strings together a set of quotations that make it appear as though patent claims should be considered “franchises,” which he contrasts with other forms of property grants that must be adjudicated in Article III court. But the term “franchise” is a synonym, not an antonym, of property. Thomas writes:

As this Court has long recognized, the grant of a patent is a matter between “the public, who are the grantors, and . . . the patentee.” [United States v.] Duell, [172 U.S. 576,] 586 [1899] (quoting Butterworth v. United States ex rel. Hoe, 112 U.S. 50, 59 (1884)). By “issuing patents,” the PTO “take[s] from the public rights of immense value, and bestow[s] them upon the patentee.” United States v. American Bell Telephone Co., 128 U.S. 315, 370 (1888). Specifically, patents are “public franchises” that the Government grants “to the inventors of new and useful improvements.” Seymour v. Osborne, 11 Wall. 516, 533 (1871) . . . . The franchise gives the patent owner “the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States.” 35 U.S.C. § 154(a)(1). That right “did not exist at common law.” Gayler v. Wilder, 10 How. 477, 494 (1851). Rather, it is a ‘creature of statute law.’ Crown Die & Tool Co. v. Nye Tool & Machine Works, 261 U.S. 24, 40 (1923).\(^\text{43}\)

The quotations employed here are worth investigating further. Do the precedents quoted shed any light on how the Court has thought about patent rights?

1. Butterworth and American Bell Telephone Co.

The quotation from Butterworth is part of larger sentence that stands for the opposite proposition:

The legislation based on this provision regards the right of property in the inventor as the medium of the public advantage derived from his invention; so that in every grant

---

36 Id.

37 Id. at 52-53.

38 59 U.S. 272.

39 For a more complete discussion, see Oil States, supra note 2, at 118-19.

40 Oil States, 138 S. Ct. at 1373 (citing Crowell, 285 U.S. at 50-51 (quoting Bakelite, 279 U.S. 438, 451 (1929)) (internal quotation marks omitted)).

41 Bakelite, 279 U.S. at 446-47 (citation omitted) (emphasis added).

42 Id. at 452.

43 Oil States, 138 S. Ct. at 1373-74.
of the limited monopoly two interests are involved, that of the public, who are the grantors, and that of the patentee.\textsuperscript{44}

The italicized words make it clear that these patents, once granted, are rights of property, which then receive the same level of protection under the Due Process Clause as all other forms of property. The same point applies to the excerpt from \textit{United States v. American Bell Telephone Co.}, which reads with full context:

In this respect the government and its officers are acting as the agents of the people, and have, under the authority of law vested in them, taken from the people this valuable privilege and conferred it as an exclusive right upon the patentee. This is property, property of a value so large that nobody has been able to estimate it.\textsuperscript{45}

\textbf{2. Seymour}

Similarly, the Court in \textit{Seymour v. Osborne} declared that “inventions secured by letters patent are property in the holder of the patent, and as such are as much entitled to protection as any other property, consisting of a franchise, during the term for which the franchise or the exclusive right is granted.”\textsuperscript{46} Qualifying “franchise” with the word “public” does nothing to change the meaning. A public franchise is a form of private property. The full passage from which Justice Thomas extracted the \textit{Seymour} quotation reads:

Letters patent are not to be regarded as monopolies, created by the executive authority at the expense and to the prejudice of all the community except the persons therein named as patentees, but as public franchises granted to the inventors of new and useful improvements for the purpose of securing to them, as such inventors, for the limited term therein mentioned, the exclusive right and liberty to make and use and vend to others to be used their own inventions, as tending to promote the progress of science and the useful arts, and as matter of compensation to the inventors for their labor, toil, and expense in making the inventions, and reducing the same to practice for the public benefit, as contemplated by the Constitution and sanctioned by the laws of Congress.\textsuperscript{47}

\textit{Seymour} rightly notes that a patent is not a monopoly but an exclusive grant that compensates for services rendered to the public at large. In attempting to rebrand patents as something less than property, Justice Thomas misreads the very precedents that prove him wrong.

\textbf{3. Gayler}

Justice Thomas next quotes \textit{Gayler v. Wilder}, which he claims stands for the proposition that “the right [to exclude others from making, using, offering for sale, or selling a patented invention in the U.S.] ‘did not exist at common law.’”\textsuperscript{48} But again, context is everything, as the full passage from which this quotation is extracted makes clear:

Now the monopoly granted to the patentee is for one entire thing—it is the exclusive right of making, using, and vending to others to be used, the improvement he has invented, and for which the patent is granted. The monopoly did not exist at common law, and the rights, therefore, which may be exercised under it cannot be regulated by the rules of the common law. It is created by the act of Congress, and no rights can be acquired in it unless authorized by statute, and in the manner the statute prescribes. . . . the patentee may assign his whole interest or an undivided part of it. But if he assigns a part under this section, it must be an undivided portion of his entire interest under the patent, placing the assignee upon an equal footing with himself for the part assigned.\textsuperscript{49}

The \textit{Gayler} litigation took place in Circuit Court, and there is no hint that any administrative forum decided, or could have decided, the case. The decision assumes that the “monopoly granted” when a patent issued is treated as a full property right. The fact that the attributes of the patent are regulated by statute instead of common law does not mean that patent rights are therefore weaker than those created at common law. Congress had required that only undivided interests could be assigned because it wanted to prevent complex interests in property rights from imposing extra burdens on licensees; \textit{Gayler} merely recognizes that questions of assignability raised difficult issues in nineteenth century law generally.\textsuperscript{50} Those difficulties carried over to the assignment of interests in patents, where additional complexities had to be addressed. None of this complexity calls into question the status of patent rights as property rights.

\textbf{4. Crown Die}

Justice Thomas similarly ignores context in \textit{Crown Die & Toll Co. v. Nye Tool & Machine Works}, which explicitly relied on \textit{Gayler} to understand the role of assignments in patent litigation.\textsuperscript{51} The action involved a breach of contract dispute that placed at issue the validity of a patent assignment. The court held that even though the parties were not diverse, jurisdiction was proper in federal court because the patent “was claimed to be valid under the statutes of the United States.”\textsuperscript{52} Justice Thomas quotes \textit{Crown Die} to make the point that patents are creatures of statute, not common law. But he fails to consider that \textit{Crown Die} treats the

\begin{itemize}
  \item \textsuperscript{44} \textit{Butterworth v. United States ex rel. Hoe}, 112 U.S. 50, 59 (1884) (emphasis added).
  \item \textsuperscript{45} \textit{United States v. American Bell Telephone Co.}, 128 U.S. 315, 370 (1888).
  \item \textsuperscript{46} 78 U.S. 516, 533 (1870).
  \item \textsuperscript{47} \textit{Id.} at 533-544.
  \item \textsuperscript{49} \textit{Oil States}, 138 S. Ct. at 1374.
  \item \textsuperscript{50} For an exhaustive discussion of the many layers of this historical difficulty in connection with personal property, see 1 Grant Gilmore, \textit{Security Interests in Personal Property} ch. 7 (1965). Patents of course were frequently analogized to real property as well. See \textit{Horne v. Dept of Agric.}, 576 U.S. ____ (2015). See \textit{Id.} at 33.
  \item \textsuperscript{51} \textit{261 U.S. 24, 40 (1923)}.
  \item \textsuperscript{52} \textit{Id.} at 33.
\end{itemize}
patent at issue as a property right that had been validly assigned under the applicable statutory rule. The case supports the proposition that a patent, once granted, is an ordinary property right entitled to full protection in an Article III court, and not a lesser right that can be administratively revoked.

5. Louisville Bridge

Unfortunately, the Oil States opinion also misreads other franchise cases. Justice Thomas writes:

Patent claims are granted subject to the qualification that the PTO has “the authority to reexamine—and perhaps cancel—a patent claim” in an inter partes review. . . . This Court has recognized that franchises can be qualified in this manner. For example, Congress can grant a franchise that permits a company to erect a toll bridge, but qualify the grant by reserving its authority to revoke or amend the franchise. See, e.g., Louisville Bridge Co. v. United States, 242 U.S. 409, 421 (1917) (collecting cases). Even after the bridge is built, the Government can exercise its reserved authority through legislation or an administrative proceeding.

This passage makes it appear as though Louisville Bridge held that any dispute over a franchise can be determined by an administrative proceeding.

This reading is wrong. The case points to the exact opposite position, namely, that a judicial proceeding is required to enforce any conditions or limitations attached to a grant. The Ohio Falls Bridge was built in 1879 to cross the Ohio River at Louisville, Kentucky. The terms of the grant to the Louisvillle Bridge Company contained a proviso “that said bridge and draws shall be so constructed as not to interrupt the navigation of the Ohio River.” During World War I, the Secretary of War gave the bridge owner notice that the bridge was out of compliance because it did not allow for sufficient horizontal clearance over the river. After a number of administrative hearings failed to resolve the dispute, the Secretary of War ordered that the bridge be repaired within three years. When the bridge owner refused to comply:

[T]he Attorney General filed a bill for an injunction in the district court; appellant answered, setting up its claims as above indicated, and the case was brought to a hearing upon stipulated facts presenting, as the sole question to be determined, the legality of the order of the Secretary of War as applied to the bridge in question.

The initial grant to the bridge owner by the government created certain property rights. No one questioned that any dispute over the scope of these rights had to take place in federal district court. The only point of contention was over the proper construction of the grant, on which Justice Mahlon Pitney (my favorite) wrote as follows:

[W]hen private rights of an indefeasible nature are sought to be derived from regulatory provisions established in the exercise of [the commerce] power, the case is peculiarly one for the application of the universal rule that grants of special franchises and privileges are to be strictly construed in favor of the public right, and nothing is to be taken as granted concerning which any reasonable doubt may be raised.

Louisville Bridge does not even hint that the United States could resolve the case—which, like the dispute in Oil States, involved property rights obtained under a grant from the government—through an administrative proceeding.

E. McCormick and Cases Cited

Whether patent disputes may be resolved in administrative proceedings is addressed in other cases, which Justice Thomas evaluates as follows:

To be sure, two of the cases make broad declarations that “[t]he only authority competent to set a patent aside, or to annul it, or to correct it for any reason whatever, is vested in the courts of the United States, and not in the department which issued the patent.” McCormick Harvesting Machine Co., supra, at 609; accord, American Bell Telephone Co., 128 U.S., at 364. But those cases were decided under the Patent Act of 1870 . . . . That version of the Patent Act did not include any provision for post-issuance administrative review. Those precedents, then, are best read as a description of the statutory scheme that existed at that time. They do not resolve Congress’ authority under the Constitution to establish a different scheme.

Once again, a fuller examination shows that Justice Thomas has misconstrued this precedent. In McCormick Harvesting Machine Co. v. Aultman, the owner of a patent was frustrated in his effort to obtain a reissue of the patent on the ground that its subject matter was not novel in light of his previous patents. Ownership of the patent then passed by assignment to the plaintiff corporation, which “abandoned the application for a reissue, and requested and obtained from the patent office the return of the original patent.”

The full passage discussing this application reads:

It has been settled by repeated decisions of this court that when a patent has received the signature of the Secretary of the Interior, countersigned by the Commissioner of

53 Oil States, 138 S. Ct. at 1375.
54 Louisville Bridge Co. v. United States, 242 U.S. 409, 414 (1917).
55 Id. at 416.
56 Id. at 417.
57 It is also worth noting that the rule of construction Justice Pitney gave favoring the public right for government land grants was not followed in patent cases. Historically, the government retained conditions over the subject matter of land grants but not grants of patents. See, e.g., Whitney v. Emmett, 29 F. Cas. 1074, 1080 (CCED Pa. 1831) (No. 17,585); see generally Adam Mossof, Who Cares What Thomas Jefferson Thought About Patents? Reevaluating the Patent “Privilege” in Historical Context, 92 CORNELL L. REV. 953, 998-1001 (2007).
58 Oil States, 138 S. Ct. at 1376.
Patents, and has had affixed to it the seal of the Patent Office, it has passed beyond the control and jurisdiction of that office, and is not subject to be revoked or cancelled by the President or any other officer of the government. United States v. Schurz, 102 U.S. 378; United States v. Am. Bell Telephone Co., 128 U.S. 315, 363. It has become the property of the patentee, and as such is entitled to the same legal protection as other property. Seymour v. Osborne, 11 Wall. 516; Campmeyer v. Newton, 94 U.S. 225; United States v. Palmer, 128 U.S. 262, 271, citing James v. Campbell, 104 U.S. 356. The only authority competent to set a patent aside, or to annul it, or to correct it for any reason whatever, is vested in the courts of the United States, and not in the department which issued the patent. Moore v. Robbins, 96 U.S. 530, 533; United States v. Am. Bell Telephone Co., 128 U.S. 315, 364; Michigan Land and Lumber Co. v. Rust, 168 U.S. 589, 593. And in this respect a patent for an invention stands in the same position and is subject to the same limitations as a patent for a grant of lands.

Each of these internal citations stands exactly for the proposition for which it is quoted in McCormick, and against the proposition for which McCormick is cited in Oil States.

United States v. Schurz involved the issuance of a land patent for which the rule on recovon was stated as follows: “But we have also held that when, by the action of these officers . . . the title to the lands has passed from the government, the question as to the real ownership of them is open in the proper courts to all considerations appropriate to the case.”61 Later, the opinion relies on Blackstone for the proposition that patents, once issued, “are then perfect grants, and no mention is made of delivery as a prerequisite to their validity. After this, they can only be revoked or annulled by scire facias or other judicial proceeding.”62 It was clear that the decision about land patents in Schurz carried over to patents for inventions, for in American Bell Telephone, the Court concluded that:

The patent, then, is not the exercise of any prerogative power or discretion by the President or by any other officer of the government, but it is the result of a course of proceeding, quasi judicial in its character, and is not subject to be repealed or revoked by the President, the Secretary of the Interior, or the Commissioner of Patents, when once issued.

Other cases follow Schurz’s lead. In Moore v. Robbins, the Supreme Court wrote:

This decision [of the commissioner] is subject to an appeal to the secretary, if taken in time. But if no such appeal be taken, and the patent issued under the seal of the United States, and signed by the President, is delivered to and accepted by the party, the title of the government passes with this delivery. With the title passes away all authority or control of the Executive Department over the land, and over the title which it has conveyed.

Michigan Land & Lumber Co. v. Rust says, to the same effect, “After the issue of the patent the matter becomes subject to inquiry only in the courts and by judicial proceedings.”65

The question in United States v. Stone was whether the United States could revoke a land patent because it had been issued without legal authority.66 The Stone Court noted that “the United States filed a bill in the Federal court of Kansas, against the Indian chiefs and Stone, to have them judicially decreed null, and the instruments themselves delivered up for cancellation.”67 Thereafter, the Court makes no bones about the legal rule:

A patent is the highest evidence of title, and is conclusive as against the government and all claiming under junior patents or titles until it is set aside or annulled by some judicial tribunal. In England this was originally done by scire facias, but a bill in chancery is found a more convenient remedy.68

Iron Silver Mining Co. v. Campbell, another land patent case, concluded that such patents are “always and ultimately a question of judicial cognizance.”69 Likewise, Noble v. Union River Logging R.R. Co. held that “[w]ith the title passes away all authority or control of the Executive Department over the land, and over the title which it has conveyed. . . . The functions of that department necessarily cease when the title has passed from the government.”70 The same rule was adopted in United States v. Missouri, Kansas & Texas Railway Co., which involved a suit in equity for the cancellation of certain land patents. As in the other

64 Moore v. Robbins, 96 U.S. 530, 533 (1878).
65 168 U.S. 589 (1897).
66 69 U.S. 525 (1864).
67 Id. at 528. For a detailed account of Stone, see Gary Lawson, Appallingly Illegal Adjudication: The AIA Through a Constitutional Lens, 41 Geo. Mason L. Rev. ___ (forthcoming 2018). Lawson develops a powerful case that the AIA’s IPR processes do not satisfy the requisites of the appointments clause insofar as it deals with inferior officers. The relevant text of the Constitution reads:

[The President] shall nominate, and by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law; but the Congress may by Law vest the Appointment of such inferior Officers, as they think proper, in the President alone, in the Courts of Law, or in the Heads of Departments.

U.S. Const. art. II, § 2, cl. 2.
68 Stone, 69 U.S. at 535.
69 135 U.S. 286, 293 (1890).
70 147 U.S. 165, 176 (1893).
cases, it was well understood that once title had passed the matter could not be decided by the executive branch.\(^71\)

**F. Patent Rights Are Property Rights**

Justice Thomas simply brushes off this uniform line of authority when he says that these precedents "are best read as a description of the statutory scheme that existed at the time."\(^72\) Unfortunately, he gets the causal inference exactly backwards. As Justice Gorsuch notes, Justice Thomas' view ignores the palpable constitutional dimension in the case: "Allowing the Executive to withdraw a patent, *McCormick* said, 'would be to deprive the applicant of his property without due process of law, and would be in fact an invasion of the judicial branch of the government by the executive.'"\(^73\) These principles apply with equal force to patents for land and patents for inventions. There was no statutory provision for IPR in the 1870 Patent Act because it was universally understood that such a provision would amount to an impermissible encroachment on the jurisdiction of Article III courts.\(^74\) Instead of confessing error on any of these issues, Justice Thomas vainly tries to sidestep Justice Gorsuch's straightforward constitutional argument by insisting that the constitutional deficit could be cured by amending the unique and anachronistic features of the 1870 Patent Act. But this just repeats his earlier mistakes and fails to explain how the entire body of nineteenth century constitutional law can be erased by a simple statutory fix that no one has ever suggested, let alone attempted.

**II. Is Inter Partes Review Good for the Patent System?**

A. Pre-Grant v. Post-Grant Review

Justice Thomas' second strand of argument is a theoretical justification for the AIA procedures governing IPR:

The primary distinction between inter partes review and the initial grant of a patent is that inter partes review occurs after the patent has issued. But that distinction does not make a difference here. Patent claims are granted subject to the qualification that the PTO has "the authority to reexamine—and perhaps cancel—a patent claim" in an inter partes review.\(^75\)

Justice Thomas offers no explanation as to why the standards for review should be the same both before and after the patent has issued. But there are strong reasons for the earlier practice that required higher post-grant standards. Before a patent is granted, it is certainly appropriate for the relevant patent authority to take steps to see that it is valid. Hence, the patentee's reliance interest increases after the issuance of a patent. Prior to issuance, the patentee knows that he proceeds to practice or license the patent at his own risk. Hence, the landscape has changed radically. The initial review process should give rise to some confidence that the patent has been rightly issued so that the odds that a post-issue review will correct some previously unknown error are lower than they would have been if the patent had initially been issued without any review at all. This argument renders problematic all iterations of post-issuance IPR.\(^76\)

B. Patent Quality and Increased Innovation

One might argue that this departure from constitutional text and practice is justified by positive empirical results showing IPR improves patent quality and spurs innovation. But the patent quality issue was nowhere mentioned in *Oil States*, and the empirical evidence on point does not show any systematic weakness in the patent system—certainly none that could not be cured by beefing up the pre-grant patent examination process.\(^77\) Indeed, adding further layers of review carries its own serious systemic risks. One such risk is that the level of scrutiny during initial review will be reduced as the likelihood of a second review increases. As Professor Dolin observes:

Congress has adopted an overly simplistic approach that can be described as "one set of eyes is good, two is better, three is better still, etc." But as it turns out, the relationship between patent quality (however defined), certainty of patent rights, and the number of levels of review is not linear. Importantly, more opportunities to challenge issued patents also means more opportunities to engage in abusive practices to undermine legitimate patent rights.\(^78\)

This conclusion should not come as a surprise. The argument for limiting post-grant review of patents is analogous to the argument for adhering to a principle of res judicata, under which a claim that has once been resolved should not be subject to a second review given the additional time, expense, and uncertainty that would necessarily result.\(^79\) That is all the more true given the dicey form of review before the PTAB, with its serious due process weaknesses.

C. Due Process

1. Reliance

In addition, the patentee's reliance interest increases after the issuance of a patent. Prior to issuance, the patentee knows that he proceeds to practice or license the patent at his own risk. Hence,

\(^71\) *Missouri, Kansas & Texas Railway Co.*, 141 U.S. 358 (1891).

\(^72\) *Oil States*, 138 S. Ct. at 1369.

\(^73\) Id. at 1385 (Gorsuch, J., dissenting) (quoting *McCormick*, 169 U.S. at 612).

\(^74\) This is the principal lesson gleaned from review of the cases cited by Justice Thomas for the opposite proposition. See *supra* sections IA-I.E.

\(^75\) *Oil States*, 138 S. Ct. at 1375 (citing *Cauzzarella*, 136 S. Ct. at 2137).

\(^76\) Furthermore, although the point was not raised in *Oil States*, it seems clear that no breach in the wall between the executive and judicial branches should be tolerated. The 1981 and 1999 statutes that introduced more limited versions of IPR should be considered just as unconstitutional as the more invasive provisions of the AIA.


\(^78\) Dolin, *supra* note 13, at 883. For an empirical investigation, see id. at Part IV, 923–31 (noting multiple challenges to key patents, and the costs of defense and delay even when the patents are found valid by IPR).

\(^79\) See *Allen v. McCurry*, 449 U.S. 90, 94 (1980) (Stewart, J.) ("As this court and other courts have often recognized, res judicata and collateral estoppel relieve parties of the costs and vexation of multiple lawsuits, conserve judicial resources, and by preventing inconsistent decisions, encourage reliance on adjudication.").
the patent applicant has to be cautious in deciding how much to invest in developing the new technology. But once the patent has been issued, the patentee rightly receives a presumption of patent validity, having gone through an exhaustive administrative process. Enabling an infringement defendant to kick a case back to the PTAB at any time will undermine the patentee’s ability to exploit the patent through use or license. For a patentee who has successfully asserted his patent multiple times in court, going back to PTAB could undercut these victories and resource expenditures. This risk is especially great if the courts are prepared after Cuozzo to defer to the PTAB’s rules on claim construction; PTAB has adopted a rule that gives claims their “broadest reasonable interpretation,” which increases the likelihood that patents will be found invalid because they cover material that is already covered by previous patents.81

2. Due Process and Separation of Powers

At this point, the question of separation of powers tends to merge with the concerns over procedural due process.82 The amalgamation of separate powers in one set of hands raises the odds of abuse by undercutting the constitutional system of checks and balances. This theme was evident in the nineteenth century cases dealing with the separation of powers; references to Blackstone’s Commentaries were not just idle embellishment. But Justice Thomas makes light of the distinctive nature of the American patent experience when he observes:

Based on the practice of the Privy Council, it was well understood at the founding that a patent system could include a practice of granting patents subject to potential cancellation in the executive proceeding of the Privy Council. The parties have cited nothing in the text or history of the Patent Clause or Article III to suggest that the Framers were not aware of this common practice.83 He is indeed correct that the Framers were aware of these common practices. But he did not look hard enough to find the abundant evidence of the differences between English and American practice that undercut his argument. As Justice Gorsuch notes in dissent:

While the Court is correct that the Constitution’s Patent Clause was written against the backdrop of English practice, it’s also true that the Clause sought to reject some of early English practice. Reflecting the growing sentiment that patents shouldn’t be used for anticompetitive monopolies over goods or businesses which had long before been enjoyed by the public, the framers wrote the Clause to protect only procompetitive invention patents that are the product of hard work and insight and add to the sum of useful knowledge.84

82 See OS II, supra note 2, at 130 (discussing Cuozzo, 136 S. Ct. at 2136, 2142-44).
83 For an argument that they are explicitly connected, see Nathan S. Chapman & Michael W. McConnell, Due Process as Separation of Powers, 121 Yale L.J. 1672, 1681-86 (2012).
84 Id. at 1385 (Gorsuch, J., dissenting) (internal quotations and citations omitted).
85 Justice Gorsuch quotes Prof. Mossoff’s work: “[A]n American patent law—has shown, the American patent system was organized by Thomas Jefferson in accordance with democratic principles at odds with the highly restrictive rules of English patents.85 Indeed, it was precisely this rejection of older English practices that led the Framers to insist on the principle of separation of powers so that the president could never claim the powers of a King. This point was made with great force in Federalist No. 69, where Hamilton writes of the president as commander-in-chief:

The President is to be commander-in-chief of the army and navy of the United States. In this respect his authority would be nominally the same with that of the king of Great Britain, but in substance much inferior to it. It would amount to nothing more than the supreme command and direction of the military and naval forces, as first General and admiral of the Confederacy; while that of the British king extends to the DECLARING of war and to the RAISING and REGULATING of fleets and armies, all which, by the Constitution under consideration, would appertain to the legislature.86

In exactly the same fashion, the King was never regarded as a role model for the United States on matters related to patents. Indeed, in one striking passage in American Bell Telephone, the Court went out of its way to point to the critical differences between the English monarch and the American president:

But whatever may have been the course of procedure usual or requisite in the English jurisprudence, to enable the king to repeal, revoke or nullify his own patents, issued under his prerogative right, it can have but little force in limiting or restricting the measures by which the government of the United States shall have a remedy for an imposition upon it or its officers in the procurement or issue of a patent. We have no king in this country; we have here no prerogative right of the crown; and letters patent, whether for inventions or for grants of land, issue not from the President but from the United States. The President has no prerogative in the matter. He has no right to issue a patent, and, though it is the custom for patents for lands to be signed by him, they are of no avail until the proper seal of the government is affixed to them.87

86 The Federalist No. 69 (Alexander Hamilton), available at http://avalon.law.yale.edu/18th_century/fed69.asp.
87 American Bell Telephone, 128 U.S. at 362-363.
The passage should stand as a warning against the concentration of power inside the executive branch, because any breakdown in the principle of separation of powers will lead to loss of the liberty which has long been part and parcel of the American constitutional tradition. It is worth repeating that the origin of the doctrine of unconstitutional conditions came in connection with the role of federal courts and by way of early decisions that no state could condition a corporate license to do business within the state on the willingness to sacrifice diversity jurisdiction. 88

The same principle limits the ability of Congress by statute to announce a rule that says henceforth all patents are issued on the condition that they will be subject to IPR. The whole point of the unconstitutional conditions doctrine is to make sure that no branch of government is able to use its unilateral powers to upset the balance of power among the several branches. There is no reason to exempt PTAB from that rule. The failure to respect separation of powers principles is an open invitation to flout the elementary conditions of procedural due process, which the PTAB does in its standard mode of operation. 89

D. PTAB v. Other Article I Courts

In this regard, it is important to contrast the serious institutional abuses inside the PTAB with the very different situation in the well-established Article I courts that deal with such matters as taxation and bankruptcy. Indeed, as a matter of first principles, I think that the independence of the judiciary would be better protected by giving judges long terms on the bench rather than life tenure, at which point the risks of entrenchment and incompetence become too large. 90 But there is no reason ever to tolerate a system in which so-called judges are appointed to administrative tribunals on a case-by-case basis, or by a head judge who is empowered to tip the balance of a case if any panel previously chosen strays from the PTAB leadership’s preferred vision.

III. Conclusion

If Oil States had been faithful to the precedents that cut overwhelmingly against its holding, the issue of bias in PTAB proceedings would not have to be raised separately under the Due Process Clause. Yet those challenges surely will arise, and sooner rather than later, at which time the Court will have to decide whether to turn a blind eye to the problem, or whether to address the internal abuse of the PTAB on a retail or a wholesale basis. The separation of powers framework should be understood as a consistent effort to nip due process violations in the bud before they infect actual cases, where, once embedded, they will be difficult to detect and to root out. But make no mistake: this issue is percolating even now, and it will present first the lower courts and then the Supreme Court with a set of challenges that they should never have to face in the first instance. But there is a high price to be paid when a clear majority of the Supreme Court signs on to an opinion that goes so wrong in its treatment of legal precedent and constitutional theory on key issues of separation of powers, due process, and unconstitutional conditions. These mistakes assume dramatic proportions in an age where the administrative state has grown too big for its breeches.

89 See discussion of Stern and Northern Pipeline, supra at section I.A.; OS I, supra note 2, at 122.
90 See Richard A. Epstein, Mandatory Retirement for Supreme Court Justices, in Reforming the Court: Term Limits for Supreme Court Justices 435 (Roger C. Cramton & Paul D. Carrington, eds., 2006).
What Happened to the Public’s Interest in Patent Law?

By Kristen Jakobsen Osenga

Note from the Editor:

This article discusses the role of the concept of the public interest in patent law, and it criticizes recent trends among judges toward using the public interest to refuse to enjoin patent infringement. 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.


Protecting intellectual property is the government’s most important tool to encourage innovation, as our country has understood since its founding.1 The Constitution provides for the grant of exclusive patent rights to “promote the progress of science and the useful arts.”2 Thomas Jefferson, who was initially skeptical of the value of patents, later remarked, “An Act of Congress authorising [sic] the issuing patents for new discoveries has given a spring to invention beyond my conception.”3 From the very first patent, issued in 1790, to the 10 millionth patent, issued in June 2018,4 the United States has seen remarkable amounts of invention and innovation largely due to its strong patent system. A strong patent system is one that effectively provides exclusive rights for invention and innovation.

The American public benefits from innovations incentivized by this patent system and relies ever more on new technologies to make life more productive, enjoyable, and comfortable. Given these benefits, one might think that the public interest in maintaining a patent system with strong incentives for inventors would be unquestioned; for a long time, it was. Recently, though, judges in patent cases have begun to erode the rights of patentees for the purported purpose of protecting the public’s interest.5 Has the public’s interest really changed? This article examines shifting interpretations of the public’s interest in patent law and explains why an accurate understanding of the public interest actually requires us to restore our strong patent system to encourage innovation.

I. The Public’s Interest in Patent Law

Patent law performs a balancing act between promoting innovation and protecting competition.6 On one hand, patents are property rights given to encourage inventors to create,

2 U.S. Const. art. 1, § 8, cl. 8.
5To be fair, there are other issues that also threaten to erode a strong U.S. patent system, such as the uncertainty surrounding patent eligible subject matter. However, this article is focused only on the use of the “public interest” to weaken patent protection.

About the Author:

Kristen Osenga is Professor of Law at the University of Richmond School of Law.
The public has an interest in innovation—that is how it accesses new and improved technology and products. To incentivize innovation, the U.S. government grants patents that give their holders the right to exclude others for a limited time from making, using, selling, offering for sale, or importing the technology covered by those patents. On the other hand, this right to exclude creates a limited monopoly, which is the antithesis of principles underlying competition law. The public also has an interest in a competitive market that produces better products at lower prices. Patent law creates “an exception to the general rule against monopolies” and balances the public’s conflicting interests in innovation and competition by granting patents only for inventions that warrant such a reward. Because of the careful balance struck by the patent system—with a high bar for patentability and a time limit on the monopoly given—as well as the importance of patents as part of a larger economic scheme, it is generally accepted that respecting patent rights is in the public interest.

Outside of the general notion of the public’s interest in an effective and reliable patent system, the topic of “public interest” is rarely discussed in patent law. The primary exception is in the imposition of remedies for patent infringement. District court judges are required to consider the public interest as a factor when deciding whether to grant an injunction against a party found to be infringing a patent. At the International Trade Commission (ITC), administrative law judges (ALJs) are statutorily required to consider the public interest before issuing an exclusion order to prevent importation of infringing goods into the United States. Although both doctrines involve the public’s interest, courts have noted that they differ due to the “long standing principle that importation is treated differently than domestic activity.”

At the stage when judges consider the public interest, the party who is facing an injunction (at the district court) or an exclusion order (at the ITC) has already been found liable for infringing a valid patent. One might assume a judge would determine that the public interest supports allowing infringement to occur rather than maintaining strong patent rights only in extraordinary cases. In the past, this has been true, but judges are increasingly invoking the public interest to deny injunctive relief. Before arguing that this shift in how judges think about the public interest is a problem that must be fixed, this article will describe the role the public interest is supposed to play at both the district courts and the ITC.

II. Public Interest at the District Courts

In patent infringement cases decided by district courts, the question of the public’s interest arises when a judge determines whether to grant an injunction that would prohibit the infringer from continuing to infringe. Historically, permanent injunctions were issued against parties found to be infringing nearly as a matter of course. The courts based this rule on the “belief that once infringement has been established denying a patentee the right to exclude is contrary to the laws of property.”

Despite this general rule in favor of injunctions, courts would very occasionally deny injunctive relief to protect the public interest. For example, in the 1930s, the Seventh Circuit denied an injunction in a case where enjoining the infringer’s use of the patented technology would leave an “entire community without any means for disposal of raw sewage.” In the 1980s, the Federal Circuit declined to issue an injunction where to do so would “cut off the supply of . . . test kits for cancer patients.” These are fairly

---

9 See, e.g., Paul Heald, Federal Intellectual Property Law & the Economics of Preemption, 76 Iowa L. Rev. 959, 973 (1991) (“The law has presumed since at least the Statute of Monopolies that only the antithesis of free competition, the grant of exclusive rights, will suffice to stimulate the optimal level of new creation.”).
11 Precision Instrument Mfg. v. Automotive Maintenance Machinery, 324 U.S. 806, 816 (1945) (“A patent by its very nature is affected with a public interest. . . . It is an exception to the general rule against monopolies and to the right to access a free and open market. The far-reaching social and economic consequences of a patent, therefore, give the public a paramount interest in seeing that patent monopolies spring from backgrounds free from fraud or other inequitable conduct and that such monopolies are kept within their legitimate scope.”). See also Pfaff v. Wells Elecs., Inc., 525 U.S. 55, 63 (1998) (“The balance between the interest in motivating innovation and enlightenment by rewarding invention with patent protection on one hand, and the interest in avoiding monopolies that unnecessarily stifle competition on the other, has been a feature of the federal patent laws since their inception.”).
12 See, e.g., Amazon.com Inc. v. Barnesandnoble.com Inc., 73 F.Supp.2d 1228, 1248-49 (W.D. Wash. 1999), vacated and remanded, 239 F.3d 1343 (Fed. Cir. 2001) (“The public has a strong interest in the enforcement of intellectual property rights. The purpose of the patent system is to reward inventors and provide incentives for further innovation by preventing others from exploiting their work. . . . Encouraging [the patent owner] to continue to innovate—and forcing competitors to come up with their own new ideas—unquestionably best serves the public interest.”).
13 See Section II., infra.
14 See Section III., infra.
15 See Spansion, Inc. v. Int’l Trade Comm’n, 629 F.3d 1331, 1359 (Fed. Cir. 2010).
19 See City of Milwaukee, 69 F.2d at 593.
20 See Hybritech, 849 F.2d at 1458.
uncontroversial examples of how the public’s interest in health and safety may outweigh the public’s interest in effective and reliable patent rights. With few exceptions, until 2006, injunctions were routinely granted unless there was a showing of strong public interest involving health and safety.21

The situation changed in 2006 when the Supreme Court, in eBay Inc. v. MercExchange, LLC, determined the Federal Circuit’s presumption in favor of issuing a permanent injunction in cases of patent infringement was in error.22 The Supreme Court instructed lower courts to instead consider a four-factor test “according to well-established principles of equity” when deciding whether to issue permanent injunctions.23 A post-eBay plaintiff seeking injunctive relief is required to show:

1. that it has suffered an irreparable injury;
2. that monetary damages are inadequate to compensate for that injury;
3. that the balance of the hardships between the plaintiff and defendant weighs in favor of the plaintiff; and
4. that the public interest would not be disserved by the injunction.24

After eBay, courts often paid lip service to the four-factor test, but continued to issue injunctions in the vast majority of cases.25

More recently, however, courts have used the discretion afforded by the eBay four-factor test to effect policy through denial of injunctive relief. For example, courts have focused on the first two factors—irreparable harm and adequate remedy—to deny injunctions to patent assertion entities.26 Patent assertion entities have been defined in various ways, but most commonly they are firms that generate income by purchasing patents and litigating against, or licensing to, other companies that are using the technology covered by the patent.27 Courts have also often found the public interest to be disserved by grant of injunctions when

the plaintiff is a patent assertion entity, although they typically rely more on the other factors.28 In other cases, courts have used the public’s interest to delay, rather than deny, injunctive relief, giving an infringer time to design around the infringed patent before being enjoined from infringing.29

Additionally, courts have begun using the four-factor test, including the public interest factor, to deny injunctive relief to companies that participate in standard setting organizations and have asserted standard essential patents (SEPs). The Federal Circuit has unequivocally stated that injunctive relief is available for infringement of SEPs, subject to the eBay four-factor test.30

In fact, the Federal Circuit notes, “the public has an interest in encouraging participation in standard-setting organizations.”31 However, courts have still sometimes held that the public interest in accessing infringing products incorporating SEP technology outweighs its interest in respecting the patentee’s property rights.32

III. Public Interest at the ITC

In the district courts, the public’s interest has been interjected via common law and the Supreme Court’s eBay decision. But the public interest is part of the ITC’s statutory scheme. As in district court, the public interest becomes important at the remedy stage, after patent infringement has been found. 19 U.S.C. § 1337(d)(1) states that:

If the Commission determines . . . that there is a violation . . ., it shall direct that the articles concerned . . . be excluded from entry into the United States, unless, after considering the effect of such exclusion upon the public health and welfare, competitive conditions in the United States economy, the production of like or directly competitive articles in the United States, and United States consumers, it finds that such articles should not be excluded from entry.33

When the statute was enacted, a Senate Committee explicitly noted that the enumerated concerns could override the exclusionary rights of the patentee.34

21 For one of the more amusing exceptions, see CF Inflight, Ltd. v. Cablecam Sys., Ltd., No. CLA 03-CV-5374, 2004 WL 234372, at 9 (E.D. Pa. Jan. 30, 2004) (in a case involving aerial photography of the Super Bowl, a judge denied a preliminary injunction in view of the public interest, stating, “While there may not exist a compelling concern for public health, there is most certainly a public demand and interest in experiencing this visual perspective.”).


23 See id.

24 Id. at 391.


26 See id. at 1988-89 (finding injunctions were granted in only 16% of cases involving patent assertion entities); Karen E. Sandrik, Reforming Patent Remedies, 67 U. MIAMI L. REV. 95, 111 (2012) (noting that patent assertion entities “are hard pressed to get an injunction” post-eBay).

27 See, e.g., Kristen Osenga, Sticks and Stones: How the FTC’s Name-Calling Misses the Complexity of Licensing-Based Business Models, 22 GEO. MASON L. REV. 1001, 1014-1016 (2015) (discussing the varying definitions given to patent assertion entities, also known as non-practicing entities or patent trolls).

28 See Seaman, supra note 25, at 1995 (finding that in 52% of the cases where an injunction was denied the court also found an injunction would preserve the public interest).

29 See, e.g., Broadcom Corp. v. Qualcomm Inc., 543 F.3d 683, 704 (Fed. Cir. 2008) (allowing a 20-month delay before providing an injunctive remedy because “an immediate permanent injunction would adversely affect the public”).

30 See Apple Inc. v. Motorola, Inc., 757 F.3d 1286, 1331-32 (Fed. Cir. 2014). A standard essential patent (SEP) is one that covers an aspect or component of a technology standard and is necessarily infringed when a standard-compliant device is made or used or when a standard-compliant service is performed.

31 See id. at 1332.

32 For example, in a case involving a number of SEPs owned by Motorola, a judge determined that the public interest required Microsoft to be able to continue its business operations because of the presence of SEPs and because Microsoft’s consumers rely on being able to use the infringing products. Microsoft Corp. v. Motorola, Inc., 871 F. Supp. 2d 1089, 1103 (W.D. Wash. 2012).


In recent years, the ITC approved new regulations which, among other things, moved the public interest to the forefront of the ITC’s analysis. A complainant must file, “concurrently with the complaint, a separate statement of public interest” explaining how the requested relief would affect public health and welfare, competitive conditions, competitive articles, and U.S. consumers.35 Respondents and others may file responses to the patentee’s public interest statement. This shift permitted additional fact-finding on matters of public interest,36 but the crux of the public interest analysis occurs when the judge decides whether to issue an exclusion order.

Although it is specifically provided for in the statute, ITC judges have rarely invoked the public interest to deny an exclusion order; injunctive relief is issued in nearly all cases in which the ITC finds patent infringement.37 In fact, in the forty years prior to 2018, the ITC determined that the public interest trumped issuance of an exclusion order in only four cases.38 Two cases from the 1980s involved fairly clear-cut issues of public health and safety. In one case, the ITC declined to issue an exclusion order “barring import of specialized hospital beds for burn victims where the domestic producer of the beds could not meet demand and there were no therapeutically comparable beds available in the U.S.”39 In another case, the ITC did not exclude acceleration tubes required for basic atomic research because the imported tubes were of a higher quality than those available from domestic suppliers and “basic scientific research . . . is precisely the kind of activity intended by Congress” when considering the public health and welfare.40 In 1997, the ITC stated expressly that, unless a case involved drugs or medical devices, it was unlikely that it would meet the public interest exception.41 Even then, the ITC has typically required not just a clear issue of health and safety, but also an inability of domestic industry to satisfy consumer demand.42

Despite its general focus on health and safety, the ITC has invoked the public interest in cases involving other concerns. For example, during an oil shortage in 1979, the ITC used the public interest to decline to exclude importation of crankpin grinders used to make components for internal combustion motors. The Commission found there was an overriding national interest in the supply of fuel-efficient automobiles in light of the oil crisis and that the domestic industry was unable to meet demand for these parts.43 Although this aspect of the public interest is broader than health and safety, it still is based in part on the inability of domestic industry to supply a product demanded by the public.

Given the decreased likelihood of obtaining injunctive relief in district court after eBay, some commentators have claimed that patentees are “flocking” to the ITC “in search of injunctions or the credible threat of injunctions.”44 Although this may have been a smart move in the past, the ITC also has started to move away from its longstanding policy of issuing injunctions except in very rare cases involving health and safety concerns where the domestic industry cannot supply enough to meet demand.

Instead, the ITC has been using the public interest to effect policy choices in the technology innovation space. Academic commentators have encouraged the ITC to do just this. For example, Colleen Chien and Mark Lemley suggested the ITC should use the discretion afforded by the required public interest inquiry to shape patent policy.45 They proposed that the ITC consider whether the value of a patentee’s technology is small compared to the value of the product of which it is a part and to allow continued infringement in cases where this is the case.46 Practitioners too have advocated the tactic of invoking the public interest at the ITC, in part because the ITC’s inability to award money damages means a denial of an exclusion order is a “total and complete victory” for infringers.47

Despite these calls to deny injunctive relief, the ITC had previously shown that it understood the public interest in an effective and reliable patent system. In 2011, the ITC issued a partial exclusion order in a case involving mobile phones using

35 See 19 C.F.R. § 210.8(b).
38 See, e.g., Riley & Allen, supra note 36, at 758-59 (2015) (“Only four ITC decisions have used the public interest exception as a means to deny an exclusion order where it was otherwise appropriate.”); Colleen V. Chien & Mark A. Lemley, Patent Holdup, the ITC, & the Public Interest, 98 CORNELL L. REV. 1, 18-19 (2012).
41 See Certain Toothbrushes and the Packaging Thereof, Inv. No. 337-TA-391, USITC Pub. 3068, at 6 (Oct. 15, 1997) (Final) (Commission Opinion on Remedy, the Public Interest, and Bonding) (“Toothbrushes are not the type of product that have in the past raised public interest concerns (such as, for example, drugs or medical devices).”); Certain Hardware Logic Emulation Systems and Components Thereof, Inv. No.
The infringer had argued that the public interest would best be served by denying an exclusion order because first responders relied on GPS systems and the ED-VO interface provided by the patented technology. Nevertheless, the Commission recognized the tension between the public’s interest in health and safety and the public’s interest in a strong patent system: “We do not accept the general proposition that, if the infringing activity is great enough, the public interest forbids a remedy.” Rather than denying an exclusion order outright, the ITC’s decision crafted a more nuanced remedy with limited exceptions to the exclusion order.

The public interest has been invoked to overrule an ITC exclusion order at higher levels within the executive branch. As part of the “smartphone patent wars” between Apple and Samsung, Samsung filed a complaint with the ITC, alleging that a number of Apple’s iPhone, iPad, and iPod Touch devices infringed Samsung’s patents. The ITC found infringement and issued an exclusion order prohibiting importation, as well as a cease-and-desist order barring sale, of the infringing devices. However, President Obama vetoed the order, claiming that the public interest counseled against this relief because Samsung’s patent was part of a technological standard and subject to fair, reasonable, and non-discriminatory licensing requirements. Commentators have argued that the executive veto was “designed as a signal to the ITC to stop issuing injunctive relief without full consideration of the public interest at stake.”

In October 2018, an ALJ at the ITC found that Apple infringed a patent owned by Qualcomm. But the judge declined to issue an exclusion order, citing the public interest. Although the full Commission has not yet weighed in on the matter, the judge’s findings regarding the public interest signal a bias against companies that participate in standard setting organizations similar to that found among district court judges. The judge noted that two suppliers are better than one when it comes to standardized technology and that, should the infringing product be excluded, the supplier would be less competitive as the technology standards progressed, which could in turn harm national security.

IV. RESTORING THE CONCEPT OF THE PUBLIC’S INTEREST IN A STRONG PATENT SYSTEM

Despite years of acknowledging that the public has a strong interest in an effective and reliable patent system, and in the technological innovations such a patent system makes possible, judges and commentators have shifted in recent years away from that perspective. It was easier to understand the courts’ and ITC’s decisions to put public health and safety ahead of patent protection in the earlier cases. After all, treating sewage and caring for burn victims certainly fall within an ordinary view of the public’s interest. But the recent shift at both the district courts and the ITC is harder to understand. These institutions are subverting traditional patent rights in the name of the public’s interest, but without fully exploring whether there really is a public interest problem at all.

The problem with the public interest analysis in these kinds of cases is two-fold. First, there is little evidence that granting injunctions would adversely affect the public’s interest. Second, the analysis neglects the interests of patent-holder plaintiffs who are actually parties to these cases. Either of these issues alone would be sufficient to require a more careful look at the public’s interest in whether injunctive relief is issued in these cases. Because both issues are generally present, it seems unlikely that the public interest would ever warrant trumping a patentee’s right to an injunction in these types of cases.

The problems to which the courts and ITC point as supporting denial of injunctive relief are at best speculative and at worst nonexistent. Consider the ALJ’s rather tenuous argument in the Qualcomm case described above: if an infringer is not allowed to continue infringing, it will be less competitive in and likely exit from a new technology area, and that will lead to national security concerns. This chain of reasoning is incredibly speculative. The development of the technology area in question, 5G mobile connectivity, is being led by numerous global companies, including Qualcomm, Intel, Samsung, Ericsson, and others, and it will be implemented and rolled out by these and countless other manufacturers. There is little evidence that any of the important players would exit the 5G space if prohibited

---

from importing infringing products, nor is there any evidence that national security concerns would result even if one of these companies did stop working in the 5G space. Decisions to override the public’s interest in a strong patent system cannot be grounded in conjecture alone.

Some judges have also based denials of injunctive relief on a doctrine called patent holdup, which has been proven to be false. Whether the plaintiff is a patent assertion entity or a participant in a standard setting organization, the concern underlying this doctrine is that the patent holder will be able to seek unfairly high licensing rates for use of their patents because of the threat of injunctive relief.⁶⁰ Although the doctrine of patent holdup has been the subject of much theoretical discussion,⁶¹ empirical research does not support it.⁶² Despite the fact that the existence and impact of patent holdup has been questioned, most judges routinely accept the theoretical concern when denying injunctive relief in these cases.⁶³ The public’s interest in an effective and reliable patent system should not be ignored in favor of a doctrine that has been shown to be false in the real world.

In addition to the public’s interest in an effective and reliable patent system, the public also has an interest in the very types of plaintiffs that have been denied injunctive relief. Patent assertion entities provide a valuable service, functioning as facilitators between inventors who cannot or do not want to manufacture their inventions and manufacturers who wish to use patented technologies.⁶⁴ Companies that participate in and submit technology innovations to standard setting organizations also provide a valuable service, allowing these organizations to arrive at the optimal technology standard for any given problem.⁶⁵

The public has an interest in the viability of both patent assertion entities and companies that participate in standard setting organizations, because both types of plaintiffs allow for more and better products to be made available on the market. In both cases, denying injunctive relief may discourage plaintiffs from continuing to participate in the field. Thus, the public has an interest not just in a strong and reliable patent system, but in a patent system that does not unduly discriminate against certain types of patent holders.

The recent shift in the patent system where district court judges and the ITC are more regularly denying injunctive relief in the name of the public interest needs to be corrected. Rather than basing the denial of injunctions and exclusion orders on speculative and tenuous reasoning or on the discredited doctrine of patent holdup, these institutions should take their mandates to consider the public interest more seriously. The public has an interest in an effective and reliable patent system. The public has an interest in more technology and innovation and a strong economy. Patent rights, including the very essence of patents—the right to exclude—need to be respected. The public’s interest depends on it.

---

⁶⁰ See, e.g., Daryl Lim, Standard Essential Patents, Troll, and the Smartphone Wars: Triangulating the End Game, 119 Penn St. L. Rev. 1, 4-7 (2014).


⁶³ There is one case where a judge rejected the infringer’s argument that patent holdup should curtail the patentee’s requested remedy, noting that the defendants “failed to present any evidence of actual hold-up.” See Ericsson Inc. v. D-Link Sys., 2013 US Dist. LEXIS 110585 (August 6, 2013), at *63-66.


Religious Liberties

Why Nineteenth Century Bans on “Sectarian” Aid Are Facialy Unconstitutional:
New Evidence on Plain Meaning

By Robert G. Natelson

Note from the Editor:

This article presents original research on the nineteenth century meaning of the word sectarian. The author argues that, based on this new evidence, bans on sectarian aid in state constitutions—often called Blaine Amendments—are likely unconstitutional on their face because they discriminate among religions.

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. As always, we also invite responses from our readers. To join the debate, please email us at info@fedsoc.org.


About the Author:

Professor of Law, The University of Montana (ret.); Senior Fellow in Constitutional Jurisprudence, the Independence Institute, Denver. The author would like to thank the Education Policy Center at the Independence Institute and its director, Pam Benigno, for their support.

In Trinity Lutheran Church of Columbia, Inc. v. Comer1 the Supreme Court struck down a Missouri state policy of restricting religious institutions from participating in grant programs. The policy arose from the state government’s efforts,2 to comply with its state constitution’s prohibition on use of public funds to benefit “any church, sect or denomination of religion.”3

Many states have prohibitions even broader than the one in the Missouri constitution. Most state constitutions adopted during the nineteenth century, unlike that currently prevailing in Missouri, identified their proscribed recipients and purposes as sectarian. This was true of Missouri’s superseded 1875 charter,3 and it is also true of charters under which many states still operate.4 For example, the current Colorado constitution, ratified in 1876, provides:

No appropriation shall be made for charitable, industrial, educational or benevolent purposes . . . to any denominational or sectarian institution or association.5

Neither the general assembly, nor any county, city, town, township, school district or other public corporation, shall ever make any appropriation, or pay from any public fund or moneys whatever, anything in aid of any church or sectarian society, or for any sectarian purpose, or to help support or sustain any school, academy, seminary, college, university or other literary or scientific institution, controlled by any church or sectarian denomination whatsoever; nor shall any

2 Mo. Const. art. I, § 7:
That no money shall ever be taken from the public treasury, directly or indirectly, in aid of any church, sect or denomination of religion, or in aid of any priest, preacher, minister or teacher thereof, as such; and that no preference shall be given to nor any discrimination made against any church, sect or creed of religion, or any form of religious faith or worship.
3 Mo. Const. (1875), art. XI, § 11:
Neither the general assembly, nor any county, city, town, township, school district, or other municipal corporation, shall ever make an appropriation, or pay from any public fund whatever, anything in aid of any religious creed, church, or sectarian purpose; or to help to support or sustain any private or public school, academy, seminary, college, university, or other institution of learning, controlled by any religious creed, church, or sectarian denomination whatever; nor shall any grant or donation of personal property or real estate ever be made by the State, or any county, city, town, or other municipal corporation, for any religious creed, church, or sectarian purpose whatever.
4 E.g., Colo. Const. (1876), art. V, § 34 & art. IX, §§ 7 & 8 (discussed infra notes 29 and 34 and accompanying text); N.D. Const. (1889), art. VII, § 5 (“No money raised for the support of the public schools of the state shall be appropriated to or used for the support of any sectarian school”); Nev. Const. (1864), art. XI, § 9 (“No sectarian instruction shall be imparted or tolerated in any school or University that may be established under this Constitution.”).
grant or donation of land, money or other personal property, ever be made by the state, or any such public corporation to any church, or for any sectarian purpose.  

No sectarian tenets or doctrines shall ever be taught in the public school . . . .  

In some cases, prohibitions against aid to sectarian organizations or for sectarian purposes were not in the state’s original constitution, but were added by amendment during the nineteenth century; many of these changes were minor alterations in wording, suggesting that no major substantive changes were contemplated.  In other cases, twentieth century constitution writers copied such prohibitions from their states’ earlier charters.  

Commentators have long argued that prohibitions against aid to sectarian groups are void under the First and Fourteenth Amendments to the U.S. Constitution. Their arguments traditionally have taken one of two forms:  

• Assuming that sectarian means merely “religious,” then a ban on aid to sectarian recipients unconstitutionally discriminates against religion in favor of non-religion. Of course, this argument is not persuasive with “strict separationist” jurists, who believe it is fully consistent with—and may be required by—the First Amendment Religion Clauses for a state to put space between its official functions and religion, so long as all religious are treated equally.

• Sectarian is principally a nineteenth century code word for “Catholic,” so the intent behind such provisions was to discriminate among religions, which almost everyone agrees is prohibited by the Religion Clauses.  

In support of the latter contention, opponents typically connect these provisions to James G. Blaine’s 1875 effort to harness anti-Catholic sentiment to his presidential ambitions by sponsoring a federal amendment barring state aid to schools controlled by any “sect” or “denomination.” Although that proposal failed, the argument goes, Blaine remained so powerful that federal territories seeking statehood felt compelled to insert anti-sectarian language in their proposed state constitutions in order to win congressional approval. In commemoration of the putative link between state constitutions and Blaine’s proposal, anti-sectarian clauses are frequently called “Blaine amendments” or “Blaine provisions.”  

However, there are some weaknesses in arguments blaming anti-Catholic sentiment or James G. Blaine for anti-sectarian provisions in state constitutions. First, the historical record does not support a link in every state between anti-Catholic animus and the state constitutional language. Second, several state

---

6 Id., art. IX, § 7 (italics added).  
7 Id. art. IX, § 8 (italics added).  
8 E.g., Nev. Const. art. XI, § 10 (added in 1880) (“No public funds of any kind or character whatever, State, County or Municipal, shall be used for sectarian purpose.”).  
9 E.g., Tex. Const. (1876), art. VII, § 5 (“nor shall the same or any part thereof ever be appropriated to or used for the support of any sectarian school”), which now reads “The permanent school fund and the available school fund may not be appropriated to or used for the support of any sectarian school.” Id., § 5(c); N.C. Const. (1866-67), art. I, § 16, which is now id., art. VII, § 11.  
10 E.g., Mt. Const. (1972), art. X, § 6:  

The legislature, counties, cities, towns, school districts, and public corporations shall not make any direct or indirect appropriation or payment from any public fund or monies, or any grant of lands or other property for any sectarian purpose or to aid any church, school, academy, seminary, college, university, or other literary or scientific institution, controlled in whole or in part by any church, sect, or denomination. (Italics added.)  


12 Cf. Zelman v. Simmons-Harris, 536 U.S. 639, 648-49 (2001) (“The Establishment Clause of the First Amendment . . . prevents a State from enacting laws that have the “purpose” or “effect” of advancing or inhibiting religion.”).  

13 E.g., Trinity Lutheran, 137 S.Ct. at 2027 (Sotomayor, J., dissenting) (hailing “this country’s longstanding commitment to a separation of church and state beneficial to both”).  

14 E.g., USCCR, supra note 11 at 11 (statement of Anthony R. Picarello, Jr.): id. at 36 (statement of Richard D. Komer); id. at 41 (“In summary, the Blaine Amendments were intended to preserve a Protestant monopoly on public education funds and to rebuff the efforts of Catholics to acquire equivalent funding for their schools . . . . ‘sectarian’ was understood to be a code word for ‘Catholic.’”). See also Richard G. Bacon, Romanism and Romer: Equal Protection and the Blaine Amendment in State Constitutions, 6 Del. L. Rev. 1, 2-5 (2003) (focusing on anti-Catholic factors as creating Blaine provisions); Jay S. Bybee & David W. Newton, Of Orphans and Vouchers: Nevada’s ‘Little Blaine Amendment’ and the Future of Religious Participation in Public Programs, 2 Nev. L.J. 551, 554-56 (2002) (same); Erica Smith, Blaine Amendments and the Unconstitutionality of Excluding Religious Options From School Choice Programs, 18 Federalist Soc’y Rev. 90 (2017) (same).  

15 Blaine’s proposal read:  

No State shall make any law respecting an establishment of religion, or prohibiting the free exercise thereof; and no money raised by taxation in any State for the support of public schools, or derived from any public fund therefor, nor any public lands devoted thereto, shall ever be under the control of any religious sect; nor shall any money so raised or lands so devoted be divided between religious sects or denominations.  

Id. at 17 (statement of K. Hollyn Hollman).  

16 Use of the word “sectarian” has been called “the watermark of a true Blaine Amendment.” Id. at 6 (statement of Anthony R. Picarello).  

17 Id. at 17 (statement of K. Hollyn Hollman) (calling the history “not uniform” and claiming that these provisions “developed independently of any bias against a particular religion”). See also id. (“the history . . . cannot be reduced to a single phenomenon”); id. at 26 (statement of Ellen Johnson) (“The history and consequences of the Blaine Amendments have little or nothing to do with anti-Catholic animus”).
anti-sectarian clauses antedate Blaine’s proposed constitutional amendment.\textsuperscript{18} Third, Blaine himself was far from anti-Catholic:

\ldots born to a Catholic mother and a father who later converted to Catholicism; as a child, he apparently was baptized in the Catholic Church. . . . He does not seem to have harbored anti-Catholic animosity, and he refused to be drawn into “any avowal of hostility or unfriendliness to Catholics.”\textsuperscript{19}

Indeed, Blaine’s amendment, although it exploited anti-Catholic animus for political support, would not have placed Catholic church schools in a position any worse than schools sponsored by other religious denominations.\textsuperscript{20}

The same evenhandedness among religions cannot be ascribed to clauses that, unlike Blaine’s amendment, specifically forbid aid for \textit{sectarian} institutions or purposes. Of course, provisions in state constitutions generally are interpreted to signify what their ratifiers understood them to mean,\textsuperscript{21} so the language means what it did when it was ratified.\textsuperscript{22} When understood in its nineteenth century context, the addition of the word \textit{sectarian} creates effects more discriminatory and sinister than anything Blaine proposed.

Part I of this article examines language from nineteenth century state constitutions to determine whether, as some claim, \textit{sectarian} meant merely “religious” or “denominational.” The texts tell us rather clearly that this was not the case—that \textit{sectarian} held a meaning quite distinct from “religious” or “denominational.” Part II surveys contemporaneous dictionary definitions and newspaper usage. Those sources show that \textit{sectarian} referred specifically to religions and religious people the speaker deemed bigoted or out of the mainstream. Part III summarizes the constitutional implications of these findings. However, this article does not discuss the standards of constitutional review or aspects of those standards such as levels of scrutiny or burdens of proof. The focus here is on the meaning of \textit{sectarian}—a subject not heretofore reported accurately in the legal literature.

\textbf{I. Nineteenth Century Constitutional Provisions Show that “Sectarian” Had a Meaning Separate from “Religious” or “Denominational”}

Although the texts of nineteenth century constitutions do not define the word \textit{sectarian}, their language and structure show that it was not merely a synonym for religious. For example, the Nebraska constitution banned “sectarian” instruction and the use of public funds for “sectarian” purposes.\textsuperscript{23} Yet the same document made it clear that public schools were to promote religion in general: “Religion, morality, and knowledge, however, being essential to good government, it shall be the duty of the legislature . . . to encourage schools and the means of instruction.”\textsuperscript{24}

In other state constitutions, “religion,” “church,” and “sectarian” appear under circumstances suggesting that the drafters were not merely stringing together synonyms. For example, the 1875 Missouri constitution provided:

Neither the general assembly, nor any county, city, town, township, school district, or other municipal corporation, shall ever make an appropriation, or pay from any public fund whatever, anything in aid of any religious creed, church, or \textit{sectarian} purpose; or to help to support or sustain any private or public school, academy, seminary, college, university, or other institution of learning, controlled by any religious creed, church, or \textit{sectarian} denomination whatever; nor shall any grant or donation of personal property or real estate ever be made by the State, or any county, city, town, or other municipal corporation, for any religious creed, church, or \textit{sectarian} purpose whatever.\textsuperscript{25}

The presumption against surplus counsels against reading the individual components of “religious creed, church, or sectarian denomination” or “religious creed, church, or sectarian purpose” as synonymous. So does the presumption that different words and phrases in the same document mean different things. Moreover, this section contains several other word lists:

\begin{itemize}
  \item “general assembly . . . county, city, town, township, school district, or other municipal corporation;”
  \item “appropriation . . . public fund;”
  \item “school, academy, seminary, college, university, or other institution of learning;”
  \item “grant or donation;” and
  \item “personal property or real estate.”
\end{itemize}

Although the meanings of some components of these lists overlap, none is simply a synonym of another. They have different meanings, thereby implying that the lists in which \textit{sectarian} appear are not to be read as repetitive.

In addition, some nineteenth century constitutions used the term \textit{sectarian} to modify “religion,” a modification that would have been unnecessary if they meant the same thing. In 1864 a

\textsuperscript{18} Id. at 14 (statement of K. Hellyn Hollman).


\textsuperscript{21} E.g., People v. Rodriguez, 112 P.3d 693, 696 (Colo. 2005).

\textsuperscript{22} \textit{Cf.} Hawke v. Smith, 252 U.S. 221, 227 (1920) (“What it meant when adopted it still means for the purpose of interpretation.”).

\textsuperscript{23} Neb. Const. (1866-67) art. VIII, § 11:

No sectarian instruction shall be allowed in any school or institution supported in whole or in part by the public funds set apart for educational purposes; nor shall the State accept any grant, conveyance, or bequest of money, lands, or other property, to be used for sectarian purposes.

\textsuperscript{24} Id., art. I, § 16.

\textsuperscript{25} Mo. Const. 1875, art. XI, sec 11; Mt. Const. (1889), art. XI, § 8 (italics added).
The legislative assembly shall encourage the promotion of intellectual, moral, scientific and agricultural improvement, by establishing a uniform system of common schools, and schools of higher grade, embracing normal, preparatory, collegiate and university departments; but no religious institution of a strictly sectarian character shall receive the aid of the state. 26

The following year another convention proposed a constitution that repeated the italicized words verbatim. 27 If sectarian meant no more than "religious," the provision would not have included the phrase "of a strictly sectarian character."

Other nineteenth century state constitutional clauses compel the same conclusion. The final Colorado Constitution (tracked closely in the 1889 Montana charter) 28—contained this language:

No religious test or qualification shall ever be required of any person as a condition of admission into any public educational institution of the state, either as a teacher or student; and no teacher or student of any such institution shall ever be required to attend or participate in any religious service whatsoever. No sectarian tenets or doctrines shall ever be taught in the public school . . . . 29

Note how the language distinguished between “religious test[s]” and “religious service[s]” for teachers as opposed to “sectarian tenets” in the material taught. Again, the change in language raises a presumption that “religious” was not the same as sectarian. Indeed, in a 1927 case interpreting this section, the Colorado Supreme Court confirmed explicitly that they were not the same. In that case, the court held that the King James Version of the Bible, while religious, was “non-sectarian” and therefore appropriate for reading in Colorado schools. 30

Similar interpretive considerations tell us that sectarian was not a synonym for “denominational.” Constitutional provisions often used sectarian in addition to versions of “denomination” as a separate concept or as a qualifier. The 1873 Pennsylvania Constitution banned aid to any “denominational or sectarian institution.” 31 The 1870 Illinois, 32 1875 Missouri, 33 and 1876 Colorado 34 constitutions did not ban aid to denominations in general but only to “sectarian denominations.” Hence, the structure and language of these instruments inform us that sectarian had a special meaning of its own.

II. THE NINETEENTH CENTURY MEANING OF "SECTARIAN"

What was that special meaning? To answer the question, I consulted nineteenth century dictionaries to learn how drafters and ratifiers of these constitutional provisions used sectarian and certain related words. I then examined contemporaneous newspapers to verify whether the dictionaries were accurately reflecting common usage.

A. The Dictionaries

The ten dictionaries I consulted were published between 1828 and 1895. Four were American, five were British, and one issued from a publisher with offices in Britain, the United States, and Australia. In chronological order of publication, the dictionaries are:

- Noah Webster, American Dictionary of the English Language (1st ed. S Converse, New York 1828) (2 vols.) [hereinafter Webster (1828)].
- Noah Webster, An American Dictionary of the English Language (3d ed. S. Converse, New York 1830) [hereinafter Webster (1830)].

To answer the question, I consulted nineteenth century dictionaries to learn how drafters and ratifiers of these constitutional provisions used sectarian and certain related words. I then examined contemporaneous newspapers to verify whether the dictionaries were accurately reflecting common usage.

B. The Newspapers

Another source of what the words meant in the nineteenth century was the newspapers of the time. I examined 19th century newspapers to verify whether the dictionaries were accurately reflecting common usage.

1. "Sectarian Test" vs. "Religious Test"

Both the Colorado and Missouri constitutions prohibited religious tests or qualifications. The Colorado Constitution prohibited any religious test or qualification, while Missouri prohibited religious or partisan tests and qualifications. The Missouri Constitution prohibited religious or partisan tests and qualifications, while the Colorado Constitution prohibited any religious test or qualification. The Missouri Constitution prohibited religious or partisan tests and qualifications, while the Colorado Constitution prohibited any religious test or qualification.

2. "Sectarian" vs. "Denominational"

Both the Colorado and Missouri constitutions prohibited denominational tests or qualifications. The Colorado Constitution prohibited any denominational test or qualification, while Missouri prohibited denominational or sectarian tests and qualifications. The Missouri Constitution prohibited denominational or sectarian tests and qualifications, while the Colorado Constitution prohibited any denominational test or qualification. The Missouri Constitution prohibited denominational or sectarian tests and qualifications, while the Colorado Constitution prohibited any denominational test or qualification.

3. "Sectarian" vs. "Religious"

Both the Colorado and Missouri constitutions prohibited religious tests or qualifications. The Colorado Constitution prohibited any religious test or qualification, while Missouri prohibited religious or partisan tests and qualifications. The Missouri Constitution prohibited religious or partisan tests and qualifications, while the Colorado Constitution prohibited any religious test or qualification. The Missouri Constitution prohibited religious or partisan tests and qualifications, while the Colorado Constitution prohibited any religious test or qualification.

4. "Sectarian" vs. "Non-sectarian"

Both the Colorado and Missouri constitutions prohibited non-sectarian tests or qualifications. The Colorado Constitution prohibited any non-sectarian test or qualification, while Missouri prohibited non-sectarian or sectarian tests and qualifications. The Missouri Constitution prohibited non-sectarian or sectarian tests and qualifications, while the Colorado Constitution prohibited any non-sectarian test or qualification. The Missouri Constitution prohibited non-sectarian or sectarian tests and qualifications, while the Colorado Constitution prohibited any non-sectarian test or qualification.
Webster (1828) defines the word this way: “prejudice,” “bigot,” or “heretic,” or (3) do both. For example, mainstream, (2) associate the word with a negative term, such as directly that a sectarian denomination in a kingdom or state.37

SECTA´RIAN, n. One of a sect;—. . . one devoted to his party; a bigot; partisan.”44

The definitions in Ogilvie’s work were as follows: Sectarian (sek-ta´ri-an), a. . . . Pertaining to a sect or sects; peculiar to a sect; strongly or bigotedly attached to the tenets and interests of a sect or religious denomination; as sectarian principles or prejudices. ‘Men of sectarian and factious spirits.’ . . .

The Encyclopaedic Dictionary’s entries are similar. It defines sectarian, as an adjective, as:

Of or pertaining to a sect or sects; strongly or bigotedly devoted to the tenets and interests of a particular sect or religious denomination; characterized by bigoted devotion

Longmuir’s Dictionary contains a list of synonyms for common words. The entry for sectarian is “see Heretic.”38 The listed synonyms for “heretic” are “schismatic, sectarian.”39 Longmuir defines “heretic” thus:

HER´E-TIC, n. One who departs from the fundamental doctrines of Christianity—Syn: Schismatic; sectarian. . . . A sectarian is one who originates or promotes a sect or distinct organization which separates from the main body of believers. Hence the expression, “a sectarian spirit,” has a slightly bad sense, which does not attach to denominational.”40

Although Longmuir defines the adjective sectarian merely as “Pertaining or peculiar to a sect,” it defines the noun sectarian as “One of a sect, or one devoted to the interest of a sect; one of a party in religion which has separated itself from the established church. See Heretic.”41 Longmuir’s entry for “catholicity” is “The faith of the early fathers and councils; freedom from sectarianism or narrowness of views.”42

The Globe Dictionary defines the adjective sectarian as “Pertaining to a sect;—devoted to a sect;—one-sided, bigoted,” and its entry for the noun is “One of a sect;—. . . one devoted to his party; a bigot; partisan.”43 The Cabinet Dictionary defines the adjective as “Pertaining or peculiar to a sect or to sects;—devoted to a sect;—hence, narrow-minded; one-sided; bigoted” and the noun as “One of a sect;—. . . one devoted to his party; a bigot; partisan.”44

35 All these dictionaries are retrievable (with some effort) from Google Books. To enable readers to examine them more conveniently, however, I have collected PDF versions of all relevant volumes online at https://i2i.org/non-legal-materials-pertaining-meaning-sectarian-19th-century-state-constitutions/.

36 2 Webster (1828) (unpaginated).

37 Webster (1830) at 735.
to a particular sect or religious denomination; peculiar to a sect. 46

It then defines sectarianism as “The quality or state of being a sectarian; the principles of sectarians; devoted adherence to a particular sect, school, or religious denomination; bigoted or partisan zeal for a particular sect.” 47 The multi-volume Century Dictionary contains this extensive entry:

. . . I. a. 1. Of or pertaining to a sect or sects; peculiar to a sect: as, sectarian interests; sectarian principles.—2. That inculcates the particular tenets of a sect: as, sectarian instruction; a sectarian book.—3. Of or pertaining to one who is bigotedly attached to a particular sect; characterized by or characteristic of bigoted attachment to a particular sect or its teachings, interests, etc.:

Zeal for some opinion, or some party, beareth out men of sectarian and factious spirits in such practices [as slander].

Barrow, Works, Sermon xviii.

The chief cause of sectarian animosity is the incapacity of most men to conceive systems in the light in which they appear to their adherents, and enter into the enthusiasm they inspire. Leeky, Europ. Morals, I. 141.

II. n. One of a sect; especially, a person who attaches excessive importance or is bigotedly attached to the tenets and interests of a sect.

But hardly less censurable, hardly less contemptible, is the tranquilly arrogant sectarian, who denies that wisdom or honesty can exist beyond the limits of his own ill-lighted chamber. Landor, Imaginary Conversations, Lucian and Timotheus.

= Syn. See heretic. 48

Webster’s Academic defines sectarian as “a. Pert[aining] to a sect, or to sects; bigotedly attached to the tenets of a denomination. — n[oun]. One of a sect. . . . Syn. — See HerETIC.” 49

Although Chambers appears to define sectarian more neutrally—“adj., pertaining to or peculiar to a sect.—n. one of a sect”—this definition depends on the following definition of “sect”: “those who dissent from an established church: those who are illiberally, attached to any opinion, or system of religion is why an editor could criticize “sectarian” influence while also mocking a proposal for dismissing religion from public life. 54

Further evidence of this non-identity appears below.

Second, sectarian had very negative associations. Newspapers frequently paired sectarian with other disparaging words: “sectarian bigotry,” “sectarian dogma,” “sectarian prejudice,” “sectarian fanaticism,” and “sectarian hatred.” An Atlanta Daily Sun story referred to “the narrow standpoint of belief; as a bigot to the Mohammedan religion; a bigot to a form of government.” 52

B. Nineteenth Century Newspapers

Nineteenth century newspapers show how these definitions worked in context. I examined two newspaper databases: (1) The New York Times collection at ProQuest Historical Newspapers and (2) the Gale Group’s Nineteenth Century U.S. Newspapers. Entering “sectarian” in the query lines generated thousands of usages amply confirming the word’s negative sense. 53 A representative sample illustrates six conclusions about the meaning of the word sectarian in the nineteenth century.

First, there is no evidence whatsoever that sectarian merely meant “religious.” The non-identity between sectarianism and religion is why an editor could criticize “sectarian” influence while also mocking a proposal for dismissing religion from public life. 54

Further evidence of this non-identity appears below.

A person who is obstinately, and unreasonably wedded to a particular religious creed, opinion, practice, or ritual. The word is sometimes used in an enlarged sense, for a person who is illiberally, attached to any opinion, or system of

52 1 Webster (1828) (unpaginated) (Emphasis added.). Other definitions did not include the reference to Islam. See, e.g., Webster’s Academic at 62:

Bigot . . . One who regards his own faith as unquestionably right, and any other as unreasonable and wicked; one blindly devoted to his own church, party, belief, or opinion.—Bigoted, a.—Bigetry [sic], n. Syn.—Prejudiced; intolerant; narrow-minded.

53 I have collected the representative examples discussed below in PDF format at https://i2i.org/non-legal-materials-pertaining-meaning-sectarian-19th-century-state-constitutions/


55 In addition to the examples in the text, see also Telegraph, Daily Rocky Mtn. News, Aug. 10, 1876, Gale Document No. GT3011717001 (“sectarian bigotry”), and the results at https://search-proquest.com, webbib.lib.umn.edu:2443/hnpnewworktimes/results/DA689F95F7464754PQ/2accountid=14593.

56 In addition to the examples in the text, see also http://find.galegroup.com, webbib.lib.umn.edu:8080/ncmp/paginate.do?tabId=7894&currentPaging=8&sort=DateDescend&src=bcrumb&inPS=true&useLimiters=true&GroupName=mlib_1_1195&prodId=NCNP&tabLimitValue=&tabLimitIndex=.

57 E.g., Canadian Department, Boston Investigator, Dec. 27, 1876, Gale Document No. GT3015847924 (“Mr. Cook strongly urged the contemplation of the above subject . . . as . . . striking at the root of sectarian dogma . . .”); see also http://find.galegroup.com/webbib.lib.umn. edu:8080/ncmp/advancedSearch.do?sessionid=7807549EB6FBA04316540FC42E8A9E.


59 E.g. Letter to the Editor, Boston Investigator, May 16, 1860, Gale Document No. GT301583153.

60 Mr. Moody in recent sermon is reported to have said . . ., Daily Rocky Mtn. News, Jan. 7, 1876, Gale Document No. GT3011719757 (“bad passions . . . sectarian hatred”).
the sectarian bigot, or that of the factious demagogue.”61 In announcing the new academic year, a professor at Colorado College assured readers that “The college had its origin, and is maintained in no narrow, exclusive or sectarian spirit.”62 A classified advertisement in a Boston newspaper coupled “Sectarian Revivals” with “Witchcraft” and other exotic phenomena.63

Third, clinging to an unpopular religion in a way incomprehensible to the majority rendered a person sectarian. A Washington, D.C. paper assailed “men, otherwise respectable for understanding and deportment, [who] are so warped by sectarian or party spirit as not to acknowledge truths as plain as axioms.”64

As a religious minority, Roman Catholics were frequent targets of “anti-sectarian” rhetoric. The New York Times ran stories about the “threat” from ‘sectarian’ Catholic Schools.65 A San Francisco paper reported a Protestant clergyman’s warnings about “sectarian” Catholics and of the risks not reading the Bible in the public schools:

Rev. Dr. Clarke . . . made a severe argument against the Roman Catholics, and asserted that the cause of this sectarian movement was that the Papacy, which was growing weak in Europe, seeks to recover its vigor on our soil. He warned the people of the Divine displeasure in seeing God’s Word banished from the school.66

In an article discussing the “sectarian question,” an editor complained that a Catholic clergyman, under cover of a state statute granting free exercise of religion, was encouraging prisoners not to attend the prison chaplain’s Protestant Sunday school. The editor urged prison authorities to prevent the priest from interfering.67 A Protestant minister wrote that readers should “rejoice in the increase of an unsectarian spirit.” But he went on to warn that if Protestants started thinking that pointing out differences among Protestant sects was “the mark of a narrow and sectarian idea of the Young Men’s Christian Association, but on a national basis, progressive and social.”68

Fourth, in contemporaneous discourse most Christians were not considered sectarian. Josiah Quincy, the president of Harvard College, explained that Unitarians, Quakers, Methodists, Baptists, Episcopalians, and the Orthodox Church were all non-sectarian. From Quincy’s remarks, an Ohio editor deduced that he thought only “Roman Catholics and the Mohammedans” were sectarian.79 Hence, a charity named for a Christian saint could be described in another article as “non-sectarian.”76

Fifth, as Quincy implicitly did, authors often contrasted sectarianism unfavorably with “good” Christianity. The New Hampshire Statesman praised a school for being “under a thoroughly Christian, not a sectarian, influence.”77 A Colorado editorial celebrated Thanksgiving by rejoicing that “What was once sectarian is now christian; that which was provincial is now national.”78 A Central City (Colorado) newspaper paper contrasted sectarian “rigidity” with Christian charity.79

Yet Catholics were not the only “sectarians.” A contributor to a Boston paper attacked “sectarian bigots” of varying denominations.69 Denver’s Rocky Mountain News referred to “Roman and other sectarian schools.”70 Mormons were tarred as sectarians.71 Among those so tarring them was President Rutherford B. Hayes.72 Some thought Jews could be sectarians.73 But Jewish speakers could turn the slur back against others. An Ohio paper reported that “A Jew proposes starting a National Young Men’s Hebrew Association, not, as he says, after the sectarian idea of the Young Men’s Christian Association, but on a national basis, progressive and social.”74

—

63 Multiple Classified Advertisements, Boston Investigator, Jan. 25, 1860, Gale Document No. GT3015812500.
66 This Afternoon’s Despatches, The Bible in Common Schools . . . The Bible in the Public Schools—The Clergy Moving in the Matter, Daily Evening Bulletin (San Francisco), March 7, 1870, Gale Document No. GT3002354333.
67 The “sectarian question” has invaded the Massachusetts State Prison at Charlestown, The Congregationalist, Boston, Mass., Dec. 6, 1876, Gale Document No. GT3004399881.
69 Letter to the Editor, Pugilistic Clergymen, Boston Investigator, Mar. 13, 1861, Gale Document No. GT3015815355.
70 At the church congress . . .”, Daily Rocky Mtn. News, Nov. 16, 1875, Gale Document No. GT3011712906 (italics added).
71 E.g., Quiet Revolutionary Movements in Mormondom, Frank Leslie’s Illustrated Newspaper, Apr. 1, 1871, Gale Document No. GT3012585419 (identifying the “Mormon system” as a “politico-sectarian concern”).
75 What is Sectarianism?, Ohio Observer, Mar. 26, 1845, Gale Document No. GT3004755960.
78 Thanksgiving Day and What It Suggests, Daily Rocky Mtn. News, Nov. 27, 1873, Gale Document No. GT3010660793; see also The Quakers, id., Dec. 5, 1875, Gale Document No. GT3011706705 (contrasting “sectarian infatuation” with “true christianity”).
79 Religious Tendency of the Times, Daily Central City Register, Jan. 17, 1872, Gale Document No. GT3016040476.
Newspapers even printed articles on how to be a good Christian and avoid sectarianism.80

Sixth, as Josiah Quincy’s list suggests, sectarian was not a mere synonym for denominational. Like Quincy, an editor observed that many colleges and universities were “organized, endowed, and fostered by leading denominations”—and were therefore “denominational” schools. But they were “not sectarian schools, like the Catholic.”81 In an article celebrating “The Denominational Spirit,” an Ohio paper quoted Reverend Dr. Skinner:

“There ought,” says Dr. S., “to be no sectarianism among Christians, notwithstanding their differences . . . No matter, I repeat, what the differences may be, the fact that they [sic] are differences among Christians is decisive that they form no sufficient basis for sectarianism.” Dr. Skinner . . . deprecated an evil sectarian spirit, as heretical and schismatic.

As an alternative to sectarian spirit, Skinner claimed, Christians should cultivate “The true denominational spirit” which “A consistent Christian will always seek and strive to bring out, in himself and in his associates.”82 These articles illustrate the difference between denominational and sectarian as the terms were used in the nineteenth century. The former was, or could be, good; the latter was always bad. Accordingly, there were good denominations and there were sectarian (bad) denominations. The difference helps explain why the Illinois, Missouri, and Colorado constitutions did not ban aid to sectarian (bad) denominations. The difference helps explain why, during the nineteenth century, “sectarian” institutions were designed to target religious groups deemed prejudiced, bigoted, or extreme. In some states, the anti-Catholic animus or other intervening events.

Under the Fourteenth Amendment’s Equal Protection Clause, state discrimination driven by animus generally is unconstitutional.87 Moreover, equal treatment of religions is at the core of the First Amendment’s Religion Clauses, applied to the states through the Fourteenth Amendment’s Due Process Clause.88 A state’s violation of this “core” equal treatment standard triggers the strict scrutiny requirement that the state demonstrate that its discrimination is narrowly tailored to advance a compelling governmental purpose.89 As suggested by the Supreme Court in Widmar v. Vincent,89 this is a very difficult standard for any state to meet.

IV. Implications for Constitutionality

State constitutional provisions adopted during the nineteenth century prohibiting aid to sectarian groups required the state to discriminate against religions that majority opinion deemed prejudiced, bigoted, or extreme. In some states, the most natural targets were Roman Catholics, but these provisions authorized discrimination against other unpopular religions as well. Because constitutional provisions are construed according to the understanding of their ratifiers,84 those provisions mean what they meant in the nineteenth century.85 Their meaning is not changed or “purged,” as some have argued,86 by easing of anti-Catholic animus or other intervening events.

Consider an analogy: Suppose a state constitution provided that “No law shall be passed impairing the freedom of speech, except for the speech of extremists.” The exception purports to enable those controlling the state government to deny speech rights to what they see as fringe groups. Therefore it facially violates the First and Fourteenth Amendments, regardless of whether the exception was directed at any particular minority at the time of adoption, or whether there was a subsequent reduction of animus toward the original target.

82 The Denominational Spirit, Ohio Observer, Jan. 9, 1850, Gale Document No. GT3004766875.
83 Supra notes 32-34 and accompanying text.
84 Supra notes 21 and 22 and accompanying text.
86 E.g., USCCR, supra note 11 at 47-48 (statement of the Anti-Defamation League).
87 E.g., Romer v. Evans, 517 U.S. 620, 632 (1996) (provision of Colorado Constitution “inexplicable by anything but animus toward the class it affects” lacks a rational basis); City of Cleburne v. Cleburne Living Ctr., 473 U.S. 432, 448 (1985) (“mere negative attitudes, or fear” cannot justify legislation targeting a particular group); cf. Palmore v. Sidoti, 466 U.S. 429, 433 (1984) (“Private biases may be outside the reach of the law, but the law cannot, directly or indirectly, give them effect.”); cf. U.S. Dept of Agric. v. Moreno, 413 U.S. 528, 534 (1973) (“bare . . . desire to harm a politically unpopular group cannot constitute a legitimate governmental interest.”).
89 E.g., City of Hialeah, 508 U.S. at 531, 546.
90 454 U.S. 263 (1981) (holding that wider separation of church and state did not meet this standard).
Modernizing the Tribal Consultation Process for Wireless Infrastructure Siting

By Jonathan S. Adelstein & J. Wade Lindsay

Note from the Editor:

This article discusses barriers to establishing wireless infrastructure, particularly with respect to tribal consultation, and proposes reforms to the current processes.

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.


About the Authors:

Jonathan S. Adelstein is the President & Chief Executive Officer of WIA—The Wireless Infrastructure Association. He served as Commissioner of the FCC from 2002 to 2009. In 2009, Mr. Adelstein was appointed to serve as the Administrator of U.S. Department of Agriculture’s Rural Utilities Service. Wade Lindsay has been a partner with Wilkinson Barker Knauer, LLP since 1998, and has over 20 years of experience in telecommunications, energy, natural resources, and administrative law.

Expeditious and efficient wireless infrastructure deployment is essential to support the wireless broadband services that are essential to the United States’ economy and future. These services support people in their work and education, support critical government services, and enable individuals to stay connected with family and friends. More than 4.6 million Americans have jobs that depend directly or indirectly on the wireless industry.1 The mobile industry overall generates more than $400 billion in total U.S. spending2 and is expected to make a value-added contribution of $1 trillion to the North American economy by 2020, representing 4.5 percent of GDP by the end of the decade.3 These trends can be expected to accelerate in the years to come with the anticipated deployment of the Internet of Things (IoT), smart communities, and next-generation 5G wireless networks. 5G technologies are expected to produce new innovation and investment across the mobile ecosystem, with unparalleled data speeds, a massive increase in IoT devices, and entirely new services and applications.4

Supporting this wireless revolution is a projected $275 billion in industry investment, which stands to inject $500 billion into the U.S. economy and create three million jobs.5 Deployment of such 5G services, however, will require wireless service providers and infrastructure developers to build much more dense wireless networks, with hundreds of thousands of new small cells, and to expand backhaul and transport facilities to provide the needed capacity and coverage.

A clear, predictable, and efficient infrastructure siting process would ensure that wireless service providers and infrastructure investment.

---

3 Press Release, GSMA, Mobile Industry to Add $1 Trillion in Value to North American Economy by 2020, Finds New GSMA Study (Nov. 1, 2016); The White House Hosts American Leadership in Emerging Technology Event, Whitehouse.gov (June 29, 2017), https://www.whitehouse.gov/blog/2017/06/29/white-house-hosts-american-leadership-emerging-technology-event (“By encouraging the advancement of emerging technologies, and by ensuring that our scientists and tech entrepreneurs can build their greatest innovations here at home, we can continue to drive American prosperity for decades to come.”).
developers are able to deploy the dense wireless networks 5G technologies require in a timely, cost-effective manner. But instead they face significant regulatory hurdles and challenges in deploying infrastructure, not only from local governments, but also in connection with the Federal Communications Commission’s (FCC or Commission) environmental and historic preservation review processes mandated by the National Environmental Policy Act (NEPA) and Section 106 of the National Historic Preservation Act (NHPA). The costs and delays inherent in these review processes are detrimental to broadband deployment, even as experience over the past few decades confirms that wireless facilities raise few environmental or historic preservation concerns. This is especially the case with respect to small wireless facilities such as those needed for 5G deployment, which are typically designed to be attached to existing structures, leaving no new environmental footprint in surrounding property and creating minimal visual impact.

One of the primary sources of the cost and delay associated with the FCC’s review under Section 106 of the NHPA is the process for the FCC’s consultation with Indian Tribes and Native Hawaiian Organizations (Tribes). There is a broad recognition in public comments filed with the FCC that the Commission’s Section 106 tribal consultation process is inefficient and requires updating. The current process enables Tribes to become de facto gatekeepers that determine when and if projects move forward. As a result, a growing number of Tribes have the power and incentive to press for exorbitant fees beyond those charged for contributing their expertise, and to expand their participation in the consultation process in order to extract additional revenue. The results are inefficiency, delay, and additional costs, none of which significantly benefits the preservation of historic sites of religious or cultural significance to Tribes.

With the leadership of FCC Chairman Ajit Pai, Commissioner Brendan Carr has been spearheading an initiative to review and overhaul the Commission’s communications infrastructure policies in three companion matters—the Wireless Infrastructure NPRM/NOI,5 Wireline Broadband Deployment NPRM/NOI,6 and the Small Cell Infrastructure PN.7 The reforms contemplated include eliminating up-front fees, clarifying the approach to tribal consultations, and adopting a clear time period for providers to deploy in cases where Tribes do not respond.8 On March 22, the Commission will vote on an order to, among other things, modernize the Section 106 tribal review process for wireless infrastructure deployments. The Commission should approve this order and establish enforceable standards and procedures that improve efficiency, accountability, and predictability for all stakeholders.

I. Tribal Consultation Under Section 106 of the NHPA

A. Identifying Wireless Facilities That May Affect Sites of Religious or Cultural Significance

Section 106 requires federal agencies to “take into account the effect of the undertaking on any historic property.”9 In carrying out this responsibility, federal agencies are required to consult with Tribes to identify any such “federal undertaking” that may affect sites of religious and cultural significance to a Tribe and to assess that effect, if any.10 This tribal consultation requirement applies regardless of whether the historic property in question is located on or off tribal lands.11 Section 106 of the NHPA acknowledges a Tribe or Tribes’ right to consult on projects located off tribal lands because those non-tribal lands may be the ancestral homelands or historical paths of a Tribe or Tribes, and thus may contain historic properties of religious and cultural significance to them.12

The FCC fulfills its Section 106 obligations with respect to wireless infrastructure by directing licensees and applicants to follow the consultation procedures developed by the Advisory Council on Historic Preservation (ACHP)13 as modified by two Nationwide Programmatic Agreements between the FCC and the ACHP that took effect in 2001 and 2005.14 The NPA

---

6 54 U.S.C. § 306108. This is despite the fact that the FCC has a statutory mandate to ensure the timely deployment of communications networks.


12 54 U.S.C. § 306108. For purposes of this article, we put aside the question of the extent to which wireless facilities siting is a “federal undertaking” for purposes of Section 106 of the NHPA. For example, the same antennas at the same locations do not go through Section 106 review if used in connection with Wi-Fi networks.

13 Id. § 306131; 36 C.F.R. § 800.2.

14 36 C.F.R. § 800.2(c)(2)(ii) (“This requirement applies regardless of the location of the historic property; see also ACHP, CONSULTATION WITH INDIAN TRIBES IN THE SECTION 106 REVIEW PROCESS: A HANDBOOK, at 8 (June 2012) (hereinafter ACHP Handbook) (“[T]he regulations require that agencies make a reasonable and good-faith effort to identify Indian tribes that may attach religious and cultural significance to historic properties that may be affected by the undertaking, even if tribes are now located a great distance away from such properties and undertakings.”).

15 ACHP Handbook at 6.

16 36 C.F.R. Part 800.

sets out a detailed Section 106 consultation process for wireless infrastructure projects. This process covers, among other things, consultation with Tribes regarding proposed wireless projects to be located on non-tribal lands. Licensees and applicants are required to use reasonable and good faith efforts to identify and contact any Tribes that may attach religious and cultural significance to historic properties that may be affected by an undertaking. 18

B. Tribal Consultation Implemented Through the NPA and the TCNS System

To facilitate this process, the NPA permits FCC applicants or licensees to use the FCC’s Tower Construction Notification System (TCNS) to notify Tribes of proposed construction within geographic areas that the Tribes have identified as potentially containing historic properties of religious and cultural significance. 19 The NPA tribal consultation process is intended to work as follows: (1) Tribes indicate in TCNS their “areas of interest,” i.e., the areas for which they would like to be notified of wireless infrastructure projects; (2) licensees or applicants enter proposed projects into the TCNS, which then notifies the Tribes that have called the proposed project locations their “areas of interest”; (3) Tribes then notify the applicants if they would like to consult on the project; 20 (4) applicants provide consulting Tribes with substantial information about the project; (5) Tribes then have an opportunity to comment to the applicants regarding whether the proposed projects may affect historic property of cultural or religious significance; and (6) the Tribes’ comments (together with the comments of other consulting parties) are included in the applicants’ final submissions to the relevant State Historic Preservation Officers (SHPO) and, where the record so warrants, the SHPOs issue findings of concurrence with proposed “no properties” or “no effect” findings. 21

This consultation process is designed to ensure that Tribes may participate as “consulting parties” in connection with the Section 106 review of wireless infrastructure projects proposed off tribal lands. As a consulting party, a Tribe has the right to identify potential sites of cultural or religious significance, advise on identification and evaluation of historic resources, comment on potential effects, and participate in the resolution of any adverse effects. 22 In other words, Tribes are entitled to have their views considered, but they do not have explicit power to block or veto a wireless infrastructure project located off tribal lands.

II. Section 106 Tribal Consultation Is Broken

A. Section 106 Tribal Consultation Impedes Wireless Infrastructure Siting

The NPA consultation process has been in operation for more than a decade, and it is clear that tribal consultation improperly and unnecessarily impedes wireless infrastructure siting. Indeed, recent public comments to the FCC demonstrate broad agreement that the FCC’s tribal consultation process imposes undue delay, costs, and burdens on wireless infrastructure projects without meaningfully promoting the preservation of sites of religious and cultural significance to Tribes. Stated broadly, these public comments show that:

- The average time required for completing the tribal consultation process is 110 days. More than 30 percent of all requests take more than 120 days to complete, 11.5 percent take more than 180 days, and 1.2 percent take more than 365 days. The longest project took more than four years. 23
- Tribes are charging fees of $1,000 to $2,000 per project before they will even engage in the Section 106 consultation process. 24 Such tribal fees rapidly become exorbitant when multiple Tribes are assessing fees on the same project. Commenters report spending at least $2,500 in tribal fees on average per site, with one spending as much as $6,300. 25

Even more troubling, however, is that these costs and burdens on wireless infrastructure siting are not serving to benefit the preservation of tribal cultural and religious resources. Evidence presented to the FCC shows that only 0.33 percent of tribal reviews of wireless infrastructure projects result in a finding that deployment of a wireless facility will have an adverse effect on historic sites of religious and cultural significance to a Tribe. 26 Other commenters noted that, based on an analysis of 17,000 infrastructure projects, more than 99 percent of Section 106 reviews resulted in a no adverse effect finding. 27


25 AARR Comments at iii.

26 Joint Association Comments at 6, 39.

27 AARR Comments at 15.
B. The Consultation Process Gives Tribes Incentive and Leverage to Delay Wireless Infrastructure Projects In Order to Secure Improper Fees

The cause of these delays and costs can be tied directly to the lack of enforceable timelines and standards in the tribal consultation process. The FCC's Section 106 process hinges on assumptions about the Tribes' role: they will reasonably define their "areas of interest"; they will respond to an initial contact made through the TCNS system within a reasonable period (i.e., 30 days); they will act in their role as consulting parties when they respond to a notification through the TCNS; and they will complete their review of a proposed facility within a reasonable time (again, 30 days). The NPA assumes that Tribes will act in a timely fashion, but it does not impose consequences if they fail to do so.

Furthermore, as noted, for facilities to be located on non-tribal lands, Tribes merely serve as "consulting parties" under Section 106, and do not have legal authority to veto wireless infrastructure projects. However, Tribes can improperly interfere with and delay project development simply by declining to act on a timely basis. By doing so, the Tribes force wireless service providers and infrastructure developers to escalate the matter to the FCC for the agency to intervene with the non-responding Tribes, adding time, cost, and uncertainty to each project. Even a single non-responding Tribe can exert a disproportionate effect on the timeline for deploying a wireless facility. For example, in one instance, it took 525 days for the applicant to complete tribal consultation for the construction of a proposed monopole because a single tribal representative for one Tribe was on extended leave and was unavailable to complete review of the proposed project. In another instance, it took 293 days for the applicant to complete tribal consultation for a collocation on an existing building, involving no ground disturbance, because two Tribes failed to timely respond. 

In fact, evidence presented to the FCC suggests that at least some Tribes view the Section 106 consultation process less as a means of protecting tribal cultural and religious resources and more as an additional source of revenue. Tribes routinely require wireless service providers and infrastructure developers to pay fees before the Tribe will respond to the TCNS notification and before any potential cultural or religious resources have been identified. At least 95 Tribes are now known to charge fees for new construction of wireless infrastructure projects on non-tribal lands. Further, "the average cost per Tribal Nation charging fees increased by 30 percent and the average fee for collocations increased by almost 50 percent between 2015 and August 2016." Indeed, fees have become an incentive for some Tribes to participate in the tribal consultation process in the first place or to expand their participation in the process. The FCC notes that "the average number of Tribal Nations notified per [wireless] project increased from eight in 2008 to 13 in August 2016 and 14 in March 2017." Further, the FCC has identified 19 Tribes that claim 10 or more states in their entirety as their "areas of interest" in TCNS, and three Tribes that claim 20 or more full states in addition to select counties as their "areas of interest." One commenter before the FCC describes a situation in which the Little Traverse Bay Band of Odawa Indians significantly expanded their areas of geographic interest in 2015 from 29 to 154 counties of interest in the TCNS, specifically to generate additional income. Comments before the FCC also describe a tribal council meeting of the Delaware Tribe of Indians at which the Tribe expressly discussed the advantages of charging fees coupled with expanding its areas of interest in TCNS. Six days after this discussion, the Tribe adopted a new fee schedule, and over the next several months, the tribal council also adopted an investment plan for the FCC fee revenue that redirected 70 percent of the fee revenue to non-Section 106 activities.

III. A Path Forward to Protect Sites of Religious and Cultural Significance to Tribes While Speeding Wireless Infrastructure Deployment

The FCC has acknowledged that the current situation is untenable. The FCC should therefore establish clear and enforceable standards coupled with better agency oversight of the tribal consultation process. Negotiated agreements and general best practices are not an adequate substitute for a concrete, enforceable process. Without establishing a finite procedural timeline, the Commission risks continuing and exacerbating the delays and concerns associated with the current tribal consultation process.

First and foremost, the FCC should resolve the delays associated with tribal consultation, including by setting a finite and enforceable timeline for completing tribal consultation. For instance, the Commission should establish specific deadlines for Tribes to respond to requests for consultation and, if a Tribe has not responded, the Tribe would have 30 days to respond or to provide written acknowledgement that it will respond. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process. The NPA assumes that Tribes will reasonably define their areas of interest, and this is how the NPA envisions implementing the Section 106 process.

---

28 See supra Section II.B.


30 See Joint Association Comments at 15.

31 Id.

32 Wireless Infrastructure NPRM/NOI, 32 FCC Rcd at 3343-44 ¶ 35.

33 Id.

34 Id.

35 Id.


not responded by the deadline, the Tribe should be deemed to have no interest, and the applicant should be permitted to proceed with the project with the understanding that it has completed the tribal consultation process with respect to that non-responding Tribe. Second, the FCC should make clear that Tribes are not authorized to charge fees for their participation in Section 106 review process generally. But the FCC should acknowledge that professional contracting services from Tribes may be appropriate in specific circumstances, and that charging fees for such professional services would be appropriate. Third, the FCC should promote transparent and efficient information sharing between Tribes and wireless service providers and infrastructure providers in order to expedite the process of identifying and differentiating between areas that do not require Section 106 review and those that may require review.

These modifications are simple, straightforward administrative matters that can be accomplished without impinging upon tribal sovereignty or limiting tribal consultation rights under Section 106 of the NHPA. Because the Section 106 tribal consultation process discussed in this article relates to projects located off tribal lands, the FCC has a general tribal trust responsibility that is fulfilled by “compliance with general regulations and statutes.” In short, the Commission has broad discretion to administer and structure the Section 106 tribal consultation process to promote predictable, efficient, and effective consultation with Tribes.

By establishing enforceable standards and providing clear guidance and oversight, the Commission can preserve the positive aspects of the Section 106 tribal consultation process while remedying the inefficiencies, delay, and additional costs that plague the process. Providing additional guidance and clarity in these ways will be a win for the Commission, Tribes, applicants, the preservation community, and the public, as an improved process will result in more rapid and efficient deployment of the wireless infrastructure and more meaningful protection of historic sites of religious and cultural significance.

39 ACHP Handbook at 2; see also Confederated Tribes of the Goshute Reservation, 177 IBLA 171, 2009 WL 1649149 (Apr. 30, 2009) (stating that BLM’s tribal trust obligations were met by complying with standard NEPA requirements); Morongo Band of Mission Indians v. FAA, 161 F.3d 569, 574 (9th Cir. 1998) (stating that, under the general trust relationship an agency should not afford a tribe “greater rights than they otherwise have under the [governing statute] and its implementing regulations”).
The GDPR: What It Really Does and How the U.S. Can Chart a Better Course

By Roslyn Layton & Julian Mcledon

Note from the Editor:

This article discusses the European Union’s new General Data Protection Rule and argues that the regulation would not work well in the United States.

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.


About the Authors:

Roslyn Layton is a visiting scholar at the American Enterprise Institute, where she focuses on evidence-based policy for information, communications, and digital technology industries. She is also a visiting researcher at Aalborg University Center for Communication, Media, and Information Technologies and a vice president at Strand Consult, both in Denmark. Julian Mcledon is a student at the University of Colorado Law School and a student member of the Federalist Society Telecommunications & Electronic Media Practice Group.

Just seven hours after the European Union’s General Data Protection Regulation (GDPR) came into effect on May 25, 2018, Austrian activist Max Schrems’ non-profit None of Your Business (NOYB) lodged four complaints with European data protection authorities (DPAs) against Google and Facebook, claiming that the platforms force users’ consent to terms of use and demanding damages of $8.8 billion. Soon after, the French advocacy group La Quadrature du Net (LQDN) filed 19 complaints, gathering support from its “Let’s attack GAFAM and their world” campaign with a declared objective to “methodically deconstruct” Google, Apple, Facebook, Amazon, and Microsoft (GAFAM) and their “allies in press and government.”

The purpose of the GDPR is to regulate the processing of personal data. The protection of persons in the processing of such data is deemed a fundamental EU right. Specifically, the GDPR is legislation from the European Parliament composed of 173 recitals which cover 45 specific regulations on data processing, 43 conditions of applicability, 35 bureaucratic obligations for businesses to perform risk assessments, hire data protection officers, and conduct international data transfers.

The European Commission’s GDPR website claims that the goals of the regulation are to give users more control of their data and to make business “benefit from a level playing field.” But the statute itself suggests another set of stakeholders: litigants, non-profit organizations, data protection professionals, and data regulatory authorities. Non-profit organizations are empowered with new rights to organize class actions, lodge complaints, and receive compensation from fines levied on firms’ annual revenue, as high as four percent of annual revenue. The 29 DPAs across the 28 member nations are charged with 35 new responsibilities to regulate data processing. While GDPR complaints against

2  Attaquons les GAFAM et leur monde, LQDN (April 17, 2018), https://www.laquadrature.net/fr/campagne_gafam.
5  GDPR, Recital 142, Article 80.
6  GDPR, Recital 141, Article 77.
7  GDPR, Recital 143, Articles 78-79, 82.
8  GDPR, Recital 143, Article 83.
leading Silicon Valley firms are noted in the press, thousands of online entities, both in the EU and abroad, have proactively shuttered their European operations for fear of getting caught in the regulatory crosshairs. Some European DPAs report that complaints have at least doubled from last year.9 Government entities find their previously unassailable social service projects now under scrutiny by users.10 In the U.S., the popular press has lauded the GDPR,11 and Senators Edward Markey, Dick Durbin, Richard Blumenthal, and Bernie Sanders have called on U.S. companies to voluntarily adopt its provisions;12 some even want to require some of the provisions.13 But a closer look at the GDPR suggests that many people misunderstand the policy, and that it creates serious and negative unintended consequences. This paper reviews those consequences considering U.S. laws and norms, urges caution about adopting GDPR-style measures, and highlights the need for careful attention in crafting any new data protection rules.

I. What Americans Need to Know About the GDPR

A. The GDPR Is About Data Protection, Not Privacy

A popular misconception about the GDPR is that it protects privacy; in fact, it is about data protection or, more correctly, data governance.14 The word “privacy” does not even appear in the final text of the GDPR, except in a footnote.15 Data privacy is about the use of data by people who are allowed to have it. Data protection, on the other hand, refers to technical systems that keep data out of the hands of people who should not have it. By its very name, the GDPR regulates the processing of personal data, not privacy.

Privacy is a complex notion having to do with being apart from others, being concealed or secluded, being free from intrusion, being let alone, and being free from publicity, scrutiny, surveillance, and unauthorized disclosure of one’s personal information.16 Data privacy is the application of these principles to information technology. The International Association of Privacy Professionals (IAPP) Glossary notes that data or information privacy is the “claim of individuals, groups or institutions to determine for themselves when, how and to what extent information about them is communicated to others.”17 Data protection, on the other hand, is the safeguarding of information from corruption, compromise, or loss. IPSwitch summarizes the difference: “data protection is essentially a technical issue, whereas data privacy is a legal one.”18 It is important to make this distinction because the terms are often used interchangeably in popular discourse but do not, in fact, mean the same thing.

Yet some assert that the GDPR is somehow a morally superior regime, conflating the high-minded value of privacy with a secular set of technical requirements on data protection.19 The Data Protection Supervisor, the new EU super-regulator for data protection, bills itself as the “global gold standard,” even though the components of the regulation that created it are relatively new and still being tested in both the marketplace and the courts.20 The GDPR itself declares in Recital 4, “The processing of personal data should be designed to serve mankind.”21 Despite EU assertions to the contrary, there are many technical forms of data protection; each has its own features, but there is no one regime which is objectively and empirically “best.”

Many Americans are persuaded by these lofty descriptions of the GDPR—contrasting them with what they see as a morally inferior laissez faire approach at home—both because they confuse data privacy and protection and because they are not familiar with America’s own substantive personal informational privacy protections developed since the founding. Journalists and commentators glibly refer to the U.S. as the “wild west,” as if there are no laws or regulation on data privacy and data protection.22 In


---

The American notion of privacy is predicated in large part on freedom from government intrusion and as a counterweight to the growth of the administrative state. The Bill of Rights’ Third, Fourth, and Fifth Amendments responded to the egregious British abuses of personal privacy, including the quartering of soldiers in private homes, the search and seizure of colonists’ property, and forcing colonists to divulge information. Some of the first laws in the new republic were enacted to protect privacy in the use of mail. These were followed by laws constraining the government’s use of the census and its ability to compel information in court. The 1966 Freedom of Information Act (FOIA) ensured that people could access records held by the government. Given this history of pushing back against government intrusion, it is reasonable to be skeptical that increasing government power is now the key to privacy in the U.S.

B. The GDPR’s Primary Goals Are Geopolitical

European leaders have expressed a positive view of the GDPR, often in terms that go beyond what it actually seeks to accomplish. But to analyze a policy like the GDPR, we must set aside the political pronouncements surrounding it and evaluate its real-world effects. Besides its effect on data processing, the GDPR can be investigated for its ability to achieve important European geopolitical goals, including (1) solidifying legitimacy for Brussels during a period of deep skepticism among voters, and (2) strengthening European political power against the real or perceived threat of American digital prowess.

The GDPR can be examined in the context of a heightened pro v. anti-EU debate, fueled by a rise in Euroscepticism and nationalist parties which charge that European integration weakens national sovereignty. Smarting from a disgruntled electorate and the Brexit bombshell, pro-European coalitions support pan-European regulation such as the GDPR to legitimize the EU project. It should be noted that Eurosceptic political actors are not necessarily opposed to data protection regulation; they merely prefer the primacy of national institutions over European ones, largely because of concerns that EU institutions and policies are subverting democracy.

In the case of the GDPR, there was no groundswell of public support calling for the enactment of greater data protection regulation. The GDPR was enacted during a period of voter “disengagement.” Participation in European Parliament elections has dwindled from 62 percent in 1979 to just 42 percent in 2014. This environment of voter disengagement is conducive for the collective action of organized special interests to defeat a diffuse, disgruntled, and unorganized majority. Relatively few Europeans are even aware of the GDPR. For example, a United Kingdom survey found that only 34 percent of respondents recognized the law, and even fewer knew what it covered. Essentially, a relatively small group of GDPR advocates successfully implemented massive pan-European regulation without significant voter buy-in. Public
have long maintained digital public databases of individual countries, with their traditions of transparency and egalitarianism, cultural differences and exigencies. For example, the Nordic protection and data privacy is underscored by demonstrated and vary considerably across countries.

C. Regulatory Approaches to Privacy Are Mediated by Cultural Norms and Vary Considerably Across Countries

The difference between U.S. and EU approaches to data protection and data privacy is underscored by demonstrated cultural differences and exigencies. For example, the Nordic countries, with their traditions of transparency and egalitarianism, have long maintained digital public databases of individual citizens’ salary and income tax records. This disclosure of financial information contradicts America’s traditions and strict laws on the protection of financial information. However, the U.S. makes criminal records available to the public at the federal, state, and county level, whereas such information is not available in the same way across the EU. Both the U.S. and European countries have had telephone books and White and Yellow Pages for decades, but had they been invented in today’s precautionary environment, it is doubtful that such valuable tools would be allowed. These differences and similarities demonstrate a key debate in the field of internet policy: the individual’s right to privacy versus the public’s right to know.

Many academic studies have documented cultural differences in opinions about privacy and their implications for policy. The existence of these cultural differences suggests that exporting the GDPR’s one-size-fits-all approach to other nations with digital platforms may not be optimal for realizing what those other countries want in terms of data protection. Consider Professor Geert Hofstede’s study of cultural dimensions of citizens of the U.S. and Germany and the potential implications for data protection. Americans score highly on individualism, geographical mobility, interacting with people they don’t know, and seeking information from others. This could explain why Americans are more comfortable with sharing information, as they anticipate benefits from doing so. Germans, in contrast, score highly on uncertainty avoidance and may be more cautious with information sharing. That the leading architects of the GDPR are German and Austrian could reflect a cultural desire to lessen or avoid what they see as uncertainty in the data-driven economy, whereas Americans may believe the benefits of sharing information in society today outweigh the risks of imperfect information about the future. These conclusions regarding the different preferences for caution when disclosing data have been noted by


Professors Robert Thomson and Steven Bellman. Furthermore, studies of privacy behavior find that it is not monolithic even within cultures. Privacy concerns can diminish with education and experience. A nation’s policy choices on data privacy and protection are imbued at least to some extent with the local and culturally relevant preferences.

The conflicting theoretical views regarding data privacy and protection are well summarized in Adam Thierer’s Permissionless Innovation: The Continuing Case for Comprehensive Technological Freedom. He describes the precautionary principle as the belief that “innovations should be curtailed or disallowed until their developers can prove they will not cause any harm to individuals, groups, specific entities, cultural norms, or various existing laws, norms or traditions,” and contrasts it with permissionless innovation, in which “experimentation with new technologies and business models should be generally permitted by default” unless a “compelling” case can be made that an innovation will bring serious harm. The EU is following the precautionary principle by enacting and enforcing the GDPR, while the U.S. subscribes to permissionless innovation by allowing innovation unless and until it has proved harmful. While the EU has deemed certain data practices presumptively harmful, it has not proved the alleged harm.

D. A Decade of GDPR-type Policy Has Not Created Greater Online Trust in the EU

The GDPR could be justified if there were evidence that the many European internet-regulation laws to date have created greater trust in the digital ecosystem, but there is no such evidence. After a decade of GDPR-type regulations—in which Europeans have endured intrusive pop-ups and disclosures on every digital property they visit—Europeans report no greater sense of trust online. As of 2017, only 22 percent of Europeans shop outside their own country (a paltry increase of 10% in a decade), demonstrating that the European Commission’s Digital Single Market goals are still elusive. Moreover, only 20 percent of EU companies are highly digitized. These are primarily large firms. Small to medium sized companies invest little to modernize and market to other EU countries. The EU has not yet offered to provide any measure that the GDPR is working to create greater trust.

A poll conducted by the U.S. Census Bureau for the National Telecommunications and Information Administration (NTIA) in 2015 and 2017 gives insight into Americans’ sense of online trust. Three-quarters of Americans report concerns about risks associated with online privacy and security, but the proportion of online households reporting privacy or security concerns “fell from 84 percent to 73 percent during this period. Similarly, the proportion of online households that said privacy concerns stopped them from doing certain online activities dropped from 45 percent to 33 percent.” The survey also notes that recent events such as the Office of Personnel Management cybersecurity breach had an impact on responders’ perception. The survey is somewhat confusing because it conflates security concerns—such as identify theft, bank fraud, and the loss of personal information—with privacy concerns. A closer look at the data reveals that Americans overall are more concerned about data security than data privacy.

The Pew Research Center surveyed expectations of online trust going forward by canvassing some 1,200 technologists, scholars, practitioners, strategic thinkers, and other leaders. They found that “48% chose the option that trust will be strengthened; 28% of these particular respondents believe that trust will stay


56 Id.


62 Id.

the same; and 24% predicted that trust will be diminished.\textsuperscript{64} Trust among Americans can also be inferred from user response to Facebook in the wake of the Cambridge Analytica revelation.\textsuperscript{65} Daily active users (DAU) on Facebook in the U.S. and Canada have held steady for the past year despite negative press coverage.\textsuperscript{66} It appears that millions of U.S. Facebook users either do not know or are not concerned about the Cambridge Analytica scandal. Overall, the platform has gained 5 million DAU since the 2016 election. This is not to say that there are not concerns about the platform. Indeed, a recent poll reports that two-thirds of Americans aged 53–72 want tech companies to be regulated like big banks, even though respondents overall have some doubts about whether governments can successfully regulate such firms.\textsuperscript{67} But it does suggest that policymakers need to be careful about generalizing about all Facebook users and adopting policies predicated on an incorrect understanding of its diverse users.

Regulatory advocates would likely describe most Facebook users as suffering from a “privacy paradox” (understanding the value of privacy but failing to practice privacy enhancing behaviors),\textsuperscript{68} but the reality may be more complex. Users interpret privacy within a context, and they don’t object to sharing information per se, only to sharing that is inappropriate based on the context.\textsuperscript{69} Many users get value from Facebook; they like having their family and friends, photo albums, and messaging all in one place. They likely understand that advertising and data collection underpin the platform and make the valuable services possible, just as advertising supported analog television, radio, and print in the past. Naturally, users expect to be treated well, but they do not necessarily expect that platform providers will never make mistakes. Indeed, users could be upset about Cambridge Analytica, but rather than quitting Facebook, they would like to see how Facebook responds to the situation by making improvements to the platform. This may be related to Facebook having a resilient “brand personality” such that users understand that it is an imperfect and evolving platform.\textsuperscript{70} Indeed, Facebook experienced an increase in engagement from U.S. users following the Cambridge Analytica revelation, as users went online to change their privacy settings.\textsuperscript{71}

However, many U.S. users do quit Facebook. Hill Holliday’s survey of Generation Z (those born since 1994) shows that so-called digital natives, who are estimated to comprise 40 percent of U.S. consumers by 2020 and of whom more than 90 percent use social media platforms, found that more than one-half had switched off social media for extended periods and one-third had canceled their social media accounts.\textsuperscript{72} Users cited time wasting as the reason for quitting twice as often as a concern about privacy. While service providers don’t like the high rates of churn on their platforms,\textsuperscript{73} they are indicative of a competitive market in which consumers find it easy to leave and try other platforms with different features.

Additionally, reports suggest that some forms of user engagement are declining.\textsuperscript{74} This could be related to Facebook changing its model to emphasize posts from family and friends over news. The most significant market response was the company losing $119 billion following its second quarter financial results, the biggest market value drop for a company on a single day in U.S. history.\textsuperscript{75} This amount is roughly 10 times the maximum fine that authorities could levy on the company under the GDPR. Moreover, Facebook’s shareholders have demanded leadership changes\textsuperscript{76} and have lodged lawsuits against the company.\textsuperscript{77} The response demonstrates that users and the marketplace can be effective regulators and is consistent with the literature about

\textsuperscript{64} Lee Rainie and Janna Anderson, \textit{The Fate of Online Trust in the Next Decade}, Pew Research Center, August 10, 2017, \url{http://www.pewinternet.org/2017/08/10/the-fate-of-online-trust-in-the-next-decade/}.


\textsuperscript{69} HELEN Nissenbaum, \textit{Privacy in Context: Technology, Policy, and the Integrity of Social Life} (2009), \url{https://www.sup.org/books/title/id=8862}.


\textsuperscript{74} Ryan Erskine, \textit{Facebook Engagement Plunging Over Last 18 Months}, \url{https://www.forbes.com/sites/ryanerskine/2018/08/15/study-facebook-engagement-sharply-plunges-over-last-18-months/60c1e575f9e4}.


corporate response to public relations disasters such as the Tylenol scare and plane crashes: firms take steps to improve safety, frequently without being compelled by government to do so.78

II. LEGAL AND POLICY RISKS OF THE GDPR

Under the GDPR, data regulators and supervisors become de facto intermediaries between consumers and firms, disrupting and distorting free exchange between the parties.79 There are significant legal and policy risks that have emerged with the GDPR, including the potential for selective enforcement, the undue empowering of litigants, and the strengthening of the European data protection bureaucracy, which is largely unaccountable to voters.80

A. Enforcement Discretion

Selective enforcement or enforcement discretion occurs when authorities choose whether and how to punish an actor which has violated the law. While selective enforcement may sometimes be more efficient, it can also produce bias, corruption, and prejudice. For example, there is evidence of bias in the selective enforcement of human rights laws,81 as well as in the selective enforcement of industrial regulation in the UK.82 A recent doctoral thesis in the European University Institute's Department of Law documents the European Commission's policy of selective law enforcement and argues that it is based upon the pillars of confidentiality, bilateralism, flexibility, and autonomy.83 While it has been pressured to increase its legitimacy by improving enforcement with standards such as transparency, trilateralism, objectivity, and accountability, the Commission has resisted, and its position has been upheld in the European Court of Justice. The thesis explains that selective enforcement is prevalent because the Commission's ability to enforce the law is limited. Indeed, the Commission is reluctant to improve standards and formalize enforcement because doing so would create administrative burdens, which would in turn decrease its efficiency.84 In an editorial blog post titled GDPR and the Abusive Potential of Selective Enforcement, the authors note that “EU bureaucrats wanted to flex their muscles against Facebook—but kicked every single European enterprise in the face with the same regulatory maneuver (the GDPR) . . . . Autocrats love rules that make everyone punishable, but don’t worry, they will practice discretion—if they like you.”85 The article notes a number of concerns about the GDPR, including that its creates a heavier compliance burden on small enterprises, which ironically strengthens large companies like Facebook; that it deters innovation by European startups; that enforcement, driven by local authorities, will create a market for “regulation avoidance”; that it requires additional corporate and government bureaucracy; and that it gives rise to “data commissars”; that it reduces the free flow of information; and that it commercializes compliance as an industry.

But GDPR architects have touted selective enforcement to reassure stakeholders, as literal interpretation of the GDPR could bring commerce to a halt. Green Party Parliamentarian Jan Philipp Albrecht,86 the “father of the GDPR,” has assured critics that GDPR investigations will not focus on small to medium enterprises, but instead “will concentrate on the bigger ones that pose a threat to many consumers.”87 He noted the firms that “already for quite a time now are under suspicion of not complying with European data protection rules” and “that have been on their screen for years will be the first to be looked at.”88 He indicated that it could be two years before cases are resolved given the process for investigation, adjudication, and appeal. If smaller companies are trying in good faith to comply with the GDPR, it would be disproportionate to sanction them, Albrecht said, noting that DPAs would more likely assist them to become compliant. The UK DPA’s Information Commissioner’s Office (ICO) noted that it will prioritize 30 named firms in its investigations under the GDPR.89 Albrecht’s comments underscore the uncertainty created by the GDPR.

B. Empowering Litigants

The GDPR also empowers litigants with a series of new rights, including rights to complain, select representatives,

84 Id.
88 Id.
and receive judicial remedy and compensation when firms fail to comply with the GDPR. The representatives of users are encouraged to create non-profit organizations to file class actions,90 lodge complaints,91 and collect recompense on behalf of users.92 These non-profits act as informal agents to surface problems and file complaints with regulatory authorities. Importantly, complaints by non-profits under the GDPR need not allege actual injury or harm—which would be required for most class actions in U.S. federal court—but only failure to comply with regulation, even if no harm results. While class actions can offer consumers a convenient, effective remedy for harm, they can also be abused by unscrupulous lawyers and by activists seeking to bypass democratic procedures.93 By legitimizing regulation by class action in the GDPR, the EU creates an incentive for legal abuse. Historically, Europe has largely eschewed “U.S.-style” class actions,94 noting that they disproportionately reward lawyers and litigation financiers over consumers.95 But policymakers have engineered the GDPR so that privacy activists can bring cases without overcoming legal barriers of standing and jurisdiction, which are traditional safeguards against the abuse of the legal system for private gain. Other problems include the ambiguity and politicization surrounding representation,96 the fact that organizations can deem themselves representatives of users without users’ consent, and a lack of clarity on whether consumers can opt out of class actions.97

Max Schrems, the Austrian activist and GDPR architect, is the face of EU data protection litigation.98 He founded NOYB, a non-profit in Austria, expressly to sue Silicon Valley companies under the GDPR. The organization’s executive board is the Who’s Who of GDPR proponents, including Albrecht, Austrian MP Josef Weidenholzer, Former European Commissioner for Fundamental Rights Paul Nemitz, the City of Vienna, the official consumer associations of Austria and Norway, and the American Electronic Privacy Information Center (EPIC).99 The board also includes Roland-Prozessfinanz, self-described as “the most experienced litigation funder in Europe,” which has financed Schrems’ lawsuits since 2014 and takes a 20-40% cut of judicial penalties levied in such cases.100

Working with Schrems is Andrea Jelinek,101 the first Chair of the European Data Protection Board (EDPB)—the pan-European data protection regulator established by the GDPR—102 and the current chief of the Austrian DPA.103 Jelinek has prioritized NOYB’s complaints against Google, Facebook, Instagram, and WhatsApp in official EU investigations and has incorporated NOYB parlance and arguments such as “forced consent” into her media talking points.104

Schrems campaigned and litigated against Facebook long before the GDPR came into effect. Following a study abroad program at California’s Santa Clara University, he launched a formal complaint against Facebook in Ireland in 2011, alleging that the company kept information he had tried to delete.105 He claimed that Facebook refused to hand over his “biometric faceprint,” saying it was a trade secret. Though Schrems eventually withdrew his complaint, the Irish Data Protection Commission audited Facebook and ordered it to delete some files and disable its facial recognition software.106 Schrems filed another complaint against Facebook in Irish court in 2013, which ultimately brought down the 15 year old safe harbor agreement between the U.S. and the EU that had facilitated $250 billion of digital trade annually.107 The following year, he launched a class action against the company in Austria and invited any Facebook user in the world to participate with a promise of “token” damages of €500 per user.108 Capped at 25,000 Facebook users, the suit is

90 GDPR, Recital 142.
91 GDPR, Recital 141.
92 GDPR, Recital 143.
93 See generally MARTIN H. REDISH, WHOLESALE JUSTICE (2009).
94 See Redish, supra note 93.
95 For example, different types of organizations can be formed to create complaints and class actions provided that they operate in the “public interest” and in the field of data protection rights. See GDPR, Chapter 8, Article 80.
96 See Rickard, supra note 94.
99 See Redish, supra note 93.
100 See Redish, supra note 93.
106 Id.
108 Lawyer Suing Facebook Overwhelmed with Support, THE GUARDIAN (Aug. 6, 2014), https://www.theguardian.com/technology/2014/aug/06/facebook-privacy-action-austria-max-schrems. This amount of proposed damages is absurd, considering that Facebook's average revenue per user in Europe was USD $13.8 in 2014. See Facebook’s average revenue per user as of 4th quarter 2017, by region (in U.S. dollars), STATISTA, https://www.
believed to be the largest class action in Europe to date.\textsuperscript{109} Austrian lower courts recognized Schrems’ financial and media interest in pursuing the case and rejected it on jurisdictional and procedural grounds.\textsuperscript{110} On two separate issues, the Austrian Supreme Court referred the case to the European Court of Justice, which ruled in January that Schrems would be allowed to proceed with the case, but that a class action was not appropriate. An appeal is pending.\textsuperscript{111}

NOYB represents the culmination and professionalization of Schrems’ effort to “confront tech giants like Facebook, Google & Co. with a team of highly qualified and motivated lawyers and IT experts on equal footing.”\textsuperscript{112} Not only is the organization positioned to serve those bringing class actions, it also serves data protection regulators, many of whom lack the training and funding to implement the GDPR.\textsuperscript{113} NOYB’s Kickstarter fundraising campaign raised €300,000 to fund operations, and its Silicon Valley-style business plan boasts that it will offer services such as group actions, collective complaints, mass mandates, abstract lawsuits, help finding favorable jurisdictions, funding of collective enforcement, a testing environment to see whether apps violate the GDPR, public relations and communications services with leading media and advocacy organizations, multilingual information, and “Statistics & Ranking to put pressure on companies and stir public debate.”\textsuperscript{114}

NOYB has several grievances about Google, Instagram, WhatsApp, and Facebook, including about their online consent processes and the use of data on platforms, but NOYB’s underlying concern is market power. NOYB claims that these companies have been so successful that they have become essential for modern life. They say that denying access to these platforms constitutes a serious detriment to people who refuse to accept new privacy policies and terms of service, so consent is essentially coerced. NOYB therefore believes that these companies should be required to provide their services even to users who do not consent to the processing of their personal data (e.g., for targeted advertising), which would effectively compel firms to provide services without compensation. According to NOYB, any consent that these companies have acquired should be viewed as invalid because users had no real choice in the matter; their options were to either consent or be kicked off an essential service. NOYB also asserts that these companies hide consent within the terms of service, making consent to processing a less than fully informed choice.\textsuperscript{115} NOYB thinks this is unfair and illegal under Recital 43 of the GDPR,\textsuperscript{116} and that essentially any processing under these conditions amounts to “exploitation of personal data.” NOYB also alleges that the platforms engage in unfair and deceptive practices to get users to consent to new terms. NOYB highlights “fake” notifications used to trick users into giving consent and requiring users to consent to new terms in order access a profile to delete it.\textsuperscript{117} NOYB urges DPAs to find that all consent to new provisions under these conditions is invalid, and it offers additional arguments in case the regulator finds that this does not constitute forced—and therefore invalid—consent. Its list of other reasons to invalidate consent includes a finding that the terms were not specific enough to permit informed consent, and that the consent was not adequately distinguishable from the privacy policy and terms of service. NOYB does not discuss whether companies can or should have any compensation for providing services, which Facebook and similar companies currently receive through targeted advertising.

The idea that Facebook has a coercive level of market power and is an essential service is contradicted by European data. Sixty-five percent of European internet users access social media of some kind,\textsuperscript{118} but Facebook’s popularity across the EU varies by country, age, gender, device, and other metrics.\textsuperscript{119} In any case, if the problem is one of market power, then antitrust is the solution, not data protection regulation. In essence, EU litigants make the same arguments as so-called “hipster” antitrust proponents in the

\begin{itemize}
\item[112] Rebecca Hill, Max Schrems Launches Privacy NGO, Wins €60k Within First 24 Hours, The Register, Nov. 29, 2017, https://www.theregister.co.uk/2017/11/29/schrems_launches_privacy_enforcement_ngo_pulls_in_nearly_60k_in_first_24_hours/.
\item[115] NOyb.Eu Filed Four Complaints, supra note 1.
\item[116] Recital 43 says: In order to ensure that consent is freely given, consent should not provide a valid legal ground for the processing of personal data in a specific case where there is a clear imbalance between the data subject and the controller, in particular where the controller is a public authority and it is therefore unlikely that consent was freely given in all the circumstances of that specific situation. Consent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations despite it being appropriate in the individual case, or if the performance of a contract, including the provision of a service, is dependent on the consent despite such consent not being necessary for such performance.
\item[117] NOyb.Eu Filed Four Complaints, supra note 1.
\item[118] See Use of Internet Services, supra note 58, at 5.
\end{itemize}
U.S.—that the evidentiary consumer welfare standard, long a transatlantic touchstone, should be abandoned in favor of the public interest, a German ordoliberal concept that the state should intervene in the market to produce a normative outcome.\(^{121}\)

Other non-profits have filed major GDPR complaints, including France’s LQDN with 19 complaints against “GAFAM” under the GDPR. LQDN’s goal is to "reverse the great farce on which GAFAM has built their world: the ‘consent’ we would give them, so that they probe our spirit and influence our wishes, is worthless.”\(^{122}\) The group vows to take action in European courts—they say “we cannot leave it entirely in the hands of the CNIL,” the French DPA—and expects its efforts to have global, not just European, consequences. LQDN argues that proper consent must be “freely” given, that it must be a positive act (i.e., there can be no pre-checked boxes indicating consent), and that it must be specific to different kinds of data processing. The complaints also argue that consent cannot be consideration for a contract to provide services. One of the LQDN’s core arguments is that developing personalized user profiles for targeted advertising is not a legitimate interest of service providers like Facebook. According to LQDN, even if the contract makes clear that such a user profile will be developed for that purpose, processing data to support advertising is not necessary for the provision of services and is therefore illegal. Some of the complaints take issue with the passive nature of consent, where check boxes come pre-ticked or ad profiles are created without the user’s explicit consent.\(^{123}\)

Additionally, LQDN has brought attention to French DPA rulings against two French startups, Teemo and Fidzup, for data protection violations.\(^{124}\) This illustrates that the French DPA has no qualms about prosecuting startups,\(^{125}\) a rebuke of the German or ad profiles are created without the user’s explicit consent.\(^{123}\) the passive nature of consent, where check boxes come pre-ticked to support advertising is not necessary for the provision of services to development personalized user profiles for targeted advertising is not a legitimate interest of service providers like Facebook. According to LQDN, even if the contract makes clear that such a user profile will be developed for that purpose, processing data to support advertising is not necessary for the provision of services and is therefore illegal. Some of the complaints take issue with the passive nature of consent, where check boxes come pre-ticked or ad profiles are created without the user’s explicit consent.\(^{123}\) Additionally, LQDN has brought attention to French DPA rulings against two French startups, Teemo and Fidzup, for data protection violations.\(^{124}\) This illustrates that the French DPA has no qualms about prosecuting startups,\(^{125}\) a rebuke of the German policymaker’s assurance that enforcement would focus on the big players.

C. Strengthening Bureaucracy

Albrecht argued that enforcement should prioritize the companies that have already been on regulators’ radar. But if the regulators already know which companies are causing problems, why require every data processor that serves Europeans to comply with preventative regulations? It could be part of a “make-work” strategy to keep Europe’s 29 DPAs in business and require firms to hire data protection officers—another GDPR requirement which is estimated to result in 75,000 new hires.\(^{126}\) From the beginning of Schrems’ lawsuits in 2011 to 2016, the Irish Data Protection Commission ballooned from 22 staff to 64, and its budget increased from $1.7 million to $5.6 million.\(^{127}\) This illustrates how activist litigation can be a form of de facto policymaking. The regulatory authority had to hire more workers to satisfy the demands of the case. This approach to staffing is different from—and less democratically accountable than—the legislative body setting up the regulator, defining its budget and mandate, and enumerating specific tasks for it to accomplish.

Those seeking to expand the role of regulatory authorities implicitly assume that those authorities have more information and therefore know better than consumers how to order transactions in the marketplace.\(^{128}\) This assumption is rarely warranted, but it is even less so where regulators do not have expertise in the area they regulate or resources to develop such expertise. The GDPR imposes massive new responsibilities on regulators without a concurrent increase in training, funding, and other resources. EU data supervisors wear many hats, including “ombudsman, auditor, consultant, educator, policy adviser, negotiator, and enforcer.”\(^{129}\) Furthermore, the GDPR widens the gap between the high expectations for data protection and the low level of skills possessed by data supervisors charged with its implementation.\(^{130}\) There are certainly many talented individuals among these ranks, but the mastery of information and communications technology varies considerably among these professionals, especially as each nation’s DPA is constituted differently.

The IAPP surveyed the complaints to the EU’s DPAs from May 25-July 31, 2018 and compared it against the same period last year.\(^{131}\) The volume, frequency, categorization, and length of complaints vary significantly across countries.\(^{132}\) Whereas Sweden, Denmark, Slovakia, Belgium, and Estonia received only a handful of complaints, others had hundreds, including the

---


123 Attaquons les GAFAM et leur monde, supra note 2.


---


Czech Republic (400), France (426), Greece (113), Ireland (933),
Netherlands (170), Poland (756), Romania (145), and Slovenia (102). A cursory review shows that the number of complaints is not
commensurate with population. Germany, the EU’s largest
country by population, did not respond to the survey. The UK
reports a doubling of complaints from the same period last
year (1,124).133 To manage the increased volume, the ICO requires
that all UK data processors pay a significant fee to cover data
protection costs, but now many government organizations are in
arrears for the payment, including the National Health Service.134

Public choice theory suggests that the EU data supervisors’
preferences are not necessarily aligned with the public interest.135
Increasing user knowledge and the quality of data protection
technology could make people better off overall, but it could
also render data protection authorities less important. While
data supervisors will not necessarily reject policies that improve
user knowledge and technology design, it is in their interest to
promote policies that increase their own resources and legitimacy
in conducting compliance and adjudication. It is notable that
the GDPR contains no discussion of efforts to improve user
education and privacy enhancing behaviors, even though these
activities are scientifically documented to improve trust in the
online ecosystem.136

III. The GDPR’s Unintended Consequences

In the months since the GDPR took effect, there have been
reports of startups closing,137 foreign news outlets pulling out of
the EU,138 the disruption of online ad markets,139 and personal
inboxes being flooded with compliance emails.140 There are related
and significant concerns about free speech, security threats,
compliance costs, and innovation deterrence.

133 Ben Chapman, Data Breach Complaints up 160% since GDPR Came into
co.uk/news/business/news/data-breach-complaints-increase-gdpr-came-
into-force-cybersecurity-a8506711.html.
134 NHS Faces Regulatory Action over Unpaid Data Protection Fees, IT PRO,
135 James C. Cooper and William E. Kovacic, Behavioral Economics:
Implications for Regulatory Behavior, SSRN Scholarly Paper, July 21,
136 See Roslyn Layton, How the GDPR Compares to Best Practices for Privacy,
so3d/papers.cfm?abstract_id=2944358.
137 Ivana Kottasová, These Companies Are Getting Killed by GDPR, CNN
(May 11, 2018), http://money.cnn.com/2018/05/11/technology/gdpr-
technology-companies-losers/index.html.
tronc.com/gdpr/latimes.com/.
139 Jessica Davies, ‘The Google Data Protection Regulation’: GDPR Is Strafing
Ad Sellers, DIGIDAY (June 4, 2018), https://digiday.com/media/google-
data-protection-regulation-gdpr-straflng-ad-sellers/.
140 Alex Hern, Most GDPR Emails Unnecessary and Some Illegal, Say Experts,
The Guardian (May 21, 2018), https://www.theguardian.com/

A. Blocked Data and Content

Since the GDPR went into effect, over 1,000 news sites
have gone dark in the EU.141 EU residents have been unable to
access Tronc Media, whose flagship newspapers include the Los
Angeles Times, the Chicago Tribune, New York Daily News, the
Hartford Courant (America’s longest running newspaper since
1764), the Orlando Sentinel, and the Baltimore Sun.142 Nor can
they access more than 60 newspapers of Lee Enterprises covering
news across 20 U.S. states.143 Blocked media is not only a problem
for the one million Americans who live in the EU who can no
longer read news and information about their hometowns, but
for Europeans who wish to learn more about the U.S. from direct
sources rather than the state-owned media, which dominate the
press and broadcasting in most EU countries.

The GDPR has affected not just American media outlets,
but also their advertisers. Given the scope of Google’s advertising
platform and its affiliates on syndicated networks, its compliance
with the GDPR has caused ripple effects in ancillary markets.
Independent ad exchanges noted prices plummeting 20 to
40 percent.144 Some advertisers report being shut out from
exchanges.145 The GDPR’s complex and arcane designations for
“controllers” and “processors” can ensnare third party chip makers,
component suppliers, and software vendors which have never
interfaced with end users, as European courts have ruled that any
part of the internet ecosystem can be liable for data breaches.146

Many American retailers, game companies, and service
providers no longer operate in the EU. The websites of Williams-
Sonoma and Pottery Barn are dark.147 The websites of scores of
other American retailers are now polluted with pop-ups and
disclosures, prompting many customers to click away. The San
Francisco-based Klout, an innovative online service that used
social media analytics to rate its users according to online social
influence, closed down completely.148 Drawbridge, a San Mateo,

141 Jeff South, More than 1,000 U.S. News Sites Are Still Unavailable in
Europe, Two Months after GDPR Took Effect, NIEMAN LAB, August 7,
2018, http://www.niemanlab.org/2018/08/more-than-1000-u-s-news-
sites-are-still-unavailable-in-europe-two-months-after-gdpr-took-effect/.
142 Alanna Petroff, LA Times takes down website in Europe as privacy rules
media/gdpr-news-websites-la-times-tronc/index.html.
143 Renae Reints, These Major U.S. News Sites Are Blocked in the EU,
blocked-gdpr/.
144 Davies, supra note 139.
145 Catherine Armitage, Life after GDPR: what next for the advertising
www.wfanet.org/news-centre/life-after-gdpr-what-next-for-the-
advertising-industry/.
146 Judgment of the Court (Grand Chamber), EU, June 5, 2018, https://
eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62016C
0210&qid=1531145885864&from=EN.
147 Amid Confusion, EU Data Privacy Law Goes into Effect, WTOPT May 25,
law-goes-into-effect/.
com/2018/05/10/rip-klout/.
California identity management company, exited the EU and sold off its ad tracking business on account of the GDPR.149 Verve, a leading mobile marketing platform with offices in six U.S. cities, closed its European operation in advance of the GDPR, impacting 15 EU employees.150 Valve, an award-winning video game company in Bellevue, Washington, shut down an entire game community rather than invest in GDPR compliance.151 Uber Entertainment, also based in Washington, similarly shut down one of its most popular games entirely after a six year run because upgrading the platform to GDPR compliance was too expensive.152 California-based Gravity Interactive no longer offers games in the EU and refunded its European customers.153 The Las Vegas-based Brent Ozar Unlimited, which offers a range of information technology and software support services, stopped serving the EU.154 San Francisco’s Payver, the dashboard camera app that pays drivers to collect road information on potholes, fallen road signs, and other inputs to build maps to improve the safety of self-driving cars, no longer supports the EU.155

Legal news website Above the Law describes the EU closures of Ragnarok Online, Unroll.me, SMNC, Tunngle, and Steel Root, noting that the GDPR is splintering the internet and that GDPR policymakers refused to listen to concerns from startups before the launch and now refuses to fix its problems.156 Even the website of the Association of National Advertisers is not available in the EU.157


155 Getpayver, Twitter, April 5, 2018, https://twitter.com/getpayver/status/981992477392437249?lang=en (“Sorry European Payver users! Come May 24th we’re discontinuing Payver support in Europe due to GDPR. Talk to your lawmakers...”).


B. Violation of U.S. Free Speech Laws and Norms

GDPR compliance is so costly and cumbersome that these entities self-censor rather than risk violating the GDPR. If the GDPR were adopted in the U.S., it would likely violate the First Amendment, as the requirements for data processing are so onerous that they would be found to limit expression. A related issue with the GDPR is the Right to Be Forgotten (RTBF), the notion that information has a finite life and that after a certain period, the information’s life is “spent” and can be deleted from the public domain. The EU asserts that the GDPR applies to data controllers anywhere in the world if they process a European citizen’s data. Similarly, RTBF proponents such as France’s CNIL attempt to force the global removal of public information in the name of data protection. For example, the French DPA has ordered Google to delete certain search results in France, and it believes that the company must therefore delete them for all countries’ search engines. Google has appealed this holding to the European Court of Justice. The European Commission, Ireland, and Greece support the company in its appeal, arguing that RTBF stretches the meaning of data protection too far.158

Indeed, the GDPR’s asserted jurisdiction outside the EU may itself be illegal—at least where the U.S. is concerned.159 The GDPR is likely unenforceable under U.S. common law, which rejects foreign rulings when they are contrary to American policy.160 The SPEECH Act, passed in 2010, supplies strong protections for First Amendment freedoms in the context of libel suits brought in foreign jurisdictions.161

C. Potentially Blocked Innovation

Many GDPR requirements are fundamentally incompatible with big data, artificial intelligence, blockchain, and machine learning, especially those that require data processors to disclose the purpose of data processing, minimize their use of data, and automate decision-making.162 For technology developers, engineers, and entrepreneurs, the GDPR creates uncertainty not only in the text of the law and its adjudication, but in that requirements and tenets of the GDPR conflict with the operation of machine learning and artificial intelligence.163 Some of the most important recent scientific advances have been the result of processing various sets of information in inventive ways—ways that neither subjects nor controllers


160 Id. at 571.

161 Id. at 572-73.


many churches have stopped printing announcements in the official register for this information. Because of the GDPR, and more. In Denmark and Sweden, these institutions retain published information on births, deaths, weddings, baptisms, potentially frustrating the operating models of mandated social toward governments and demand erasure of their data from disease and breast cancer. 167 While many regulatory advocates world, which promises groundbreaking therapies for Alzheimer’s warehouse, the oldest and most complete genetic record in the public health databases to make advances in medicine. However, a privacy panic is threatening to derail some projects, including Iceland’s genome warehouse, the oldest and most complete genetic record in the world, which promises groundbreaking therapies for Alzheimer’s disease and breast cancer. 167 While many regulatory advocates focus attention on Silicon Valley firms and call for greater regulation, their campaign is backfiring as users turn their ire toward governments and demand erasure of their data from national health care records and other government services, potentially frustrating the operating models of mandated social programs. 168 With the mantra of “if in doubt, opt out,” about half a million Australians rejected the country’s national electronic health record, causing the computer system to crash in July 2018. 169

For centuries, European state churches have collected and published information on births, deaths, weddings, baptisms, and more. In Denmark and Sweden, these institutions retain the official register for this information. Because of the GDPR, many churches have stopped printing announcements in the bulletins for their local congregations unless they obtain consent first. 170 GDPR risks have also been identified with respect to convicted felons successfully removing information about their crimes from search engines, 172 the exchange of business cards, 172 the taking of pictures in public, 173 and disclosures of health and injury information in the trade of soccer players. 174

D. Security Concerns

A key unintended consequence of the GDPR is that it undermines the transparency of the international systems and architecture that organize the internet. The WHOIS query and response protocol for internet domain names, IP addresses, and autonomous systems is used by law enforcement, cybersecurity professionals and researchers, and trademark and intellectual property rights holders. 175 The Internet Corporation for Assigned Names and Numbers (ICANN) recently announced a Temporary Specification that allows registries and registrars to obscure WHOIS information they were previously required to make public, ostensibly in order to comply with the GDPR. 176 This could hinder efforts to combat unlawful activity online, including identity theft, cyber-attacks, online espionage, theft of intellectual property, fraud, unlawful sale of drugs, human trafficking, and other criminal behavior, and it is not even required by the GDPR.

The GDPR does not apply at all to non-personal information and states that disclosure of even personal information can be warranted for matters such as consumer protection, public

165 Id.
168 Howell, supra note 10.
169 Id.

172 Thomas Ildskov, Mundskav! Derfor holdes omfanget af FCK-spillers skade hemmelig, B.T., July 12, 2018, https://www.bt.dk/content/item/1197424.
safety, law enforcement, enforcement of rights, cybersecurity, and combating fraud. Moreover, the GDPR does not apply to domain names registered to U.S. registrants by American registrars and registries. Nor does it apply to domain name registrants that are companies, businesses, or other legal entities, rather than “natural persons.” All the same, actors including ICANN are practicing voluntary censorship because the GDPR’s provisions are so vague and the potential penalties so high. GDPR proponents have likely contributed to the impression that the GDPR urges measures like the Temporary Specification. For example, in her role in the Article 29 Working Party, the group that drove the promulgation of the GDPR, Jelinek said that the elimination and masking of WHOIS information is justified under the GDPR.177

The WHOIS problem can be described as the conflict between the individual’s right to privacy and the public’s right to know.178 It can also be understood within the context of the problem of “privacy overreach,”179 in which the drive to protect privacy becomes absolute, lacks balance with other rights, and unwittingly brings worse outcomes for privacy and data protection.180 The situation harkens back to a key fallacy of privacy and data protection law is significantly faster than that of other kinds of law, leading one scholar to suggest that it threatens to upend the balance with other fundamental rights.181 This point is eloquently underscored by Richard Epstein in his critique of the idea of privacy rights established by the Warren Court. This Progressive theory assumes that it is “always easy, if not inevitable, to expand the set of rights without adverse social consequences,” but never stops to consider that, when rights are expanded, correlative duties are imposed on others.182

E. Compliance Costs

To do business in the EU and comply with the GDPR, firms with 500 employees or more will likely have to spend between $1 and $10 million each.183 With over 19,000 U.S. firms of this size,184 total GDPR compliance costs for U.S. firms alone could reach $150 billion, twice what the U.S. spends on network investment185 and one-third of annual e-commerce revenue in the U.S.186 Economist Hosuk Lee-Makiyama calculates that the GDPR’s requirements on cross-border trade flows will increase prices, amounting to a direct welfare loss of €260 per European citizen.187 The net effect is that those companies that can afford to comply will do so, and the rest will exit. Hence the GDPR will become a barrier to market entry, punishing small firms, rewarding the largest players, and creating a codependent relationship between regulators and the firms they regulate. This is a perverse outcome for a regulation that promised to level the playing field on data protection.

IV. Conclusion

Many American policymakers have wisely recognized that the GDPR is not appropriate for the U.S. However, they are seeking to review and update existing information privacy laws to ensure consistency and effectiveness while avoiding fragmentation with state level rules.188 The Trump administration has tasked the NTIA to develop—through public comment and scientific inquiry—a set of principles that will provide a high level of protection for individuals while giving organizations


181 See Hurwitz and Jaffer, supra note 179.

182 See Brkan, supra note 180.
legal clarity and the flexibility to innovate. These principles can inform bipartisan efforts for consumer online privacy legislation under consideration in Congress. The scientific research on data protection and privacy suggests that consumer education and privacy enhancing technologies are essential to creating trust online, but these inputs are ignored in both the GDPR and the California Consumer Privacy Act. Congress can foster the continued prosperity of the information economy by ensuring that consumers can access privacy education to make informed choices, that safe harbors for privacy-enhancing innovation protect the testing and learning of new technologies, and that common standards for competition and consumer protection online are equally guaranteed for all Americans and delivered by the FTC.

This paper has reviewed perspectives on the GDPR, misconceptions about the policy, legal risks, and the GDPR's unintended consequences. The purpose of the GDPR is not to protect privacy, but rather to regulate data processing. In the past decade, the increasing data protection rules have not resulted in improved trust or increased cross-border commerce in the EU. The likelihood of selective enforcement of the GDPR, the empowerment of litigants to bring class action lawsuits, and the strengthening of the EU administrative state all suggest that the GDPR is more than the humanitarian effort it purports to be. The GDPR's unintended consequences include violations of the freedom of speech, closures of startups, blocked foreign news outlets, the disruption of online ad markets, the compromising of the WHOIS database, and the hampering of innovation. These are important realities which U.S. policymakers should consider as they evaluate whether and how to regulate online data in the U.S.


193 Specific recommendations can be found in Layton, supra note 191.
Kimberly Robinson's *The Supreme Court in Crisis* is the latest from the University of Pennsylvania’s *American Justice* series, an annual review of Supreme Court cases and happenings published shortly after a term ends. This breezy, informative account of the 2016 Term is largely even-handed and particularly useful for the reader wanting a quick overview of the term.

The book is organized by chapters discussing different themes from the term identified by Robinson. Robinson's challenge is that, as she notes, the term was characterized by a “light docket with relatively inconsequential cases.” This resulted from the fact that for the first six of the nine months it was in session, the Court had only eight justices, which, Robinson writes, “impacted not only the kinds of cases the justices took but also the way that they resolved them.” The Court chose (or was forced by circumstances) to avoid the most difficult, closely-contested cases until a new justice could be confirmed, which Robinson believes “cast it even more as a political institution.”

In the chapter “Stand Idly By,” Robinson reviews controversial cases that the Court delayed deciding or sidestepped altogether. Most notable is *Trinity Lutheran Church of Columbia v. Comer*, for which certiorari was granted a month before Justice Scalia passed in February 2016, but oral argument not held until April 19, 2017—nine days after Neil Gorsuch was sworn in. Robinson also points to the Court declining to review cases involving changes to voting requirements in Texas and North Carolina in the wake of 2013’s Voting Rights Act decision, *Shelby County v. Holder*. Chief Justice Roberts took the unusual step of explaining why the Court denied certiorari in those cases, which Robinson interprets as protesting too much in anticipation of criticism that the Court was trying to avoid difficult decisions.

In a chapter entitled “Quarter Loaf Outcomes,” Robinson writes that another tactic the Court used to deal with the incomplete complement of justices was to decide certain cases on the narrowest possible grounds. As an example, she cites *Salman v. United States*, which held that disclosing confidential company information to a close relative as a gift could constitute insider trading, but which gave no guidance as to whether such gifts to non-relatives would. Similarly, in determining whether Miami had standing to sue Bank of America for discriminatory lending practices, the Court only decided that the city could constitute an “aggrieved party” if neighborhoods blighted by widespread foreclosures resulted in increased expenditures for municipal services; to the dismay of some observers, however, the Court left another prong of the standing analysis—whether the city’s increased expenses were proximately caused by discriminatory lending, or by something else—to be decided by the lower court on remand.

Due in large part to its cautious approach, the 2016 Term was marked by an unusually high degree of consensus and a greater-than-usual number of unanimous decisions, with only two dissents read from the bench. Along with the absence of divisive, blockbuster cases, the relative consensus among the justices makes

---

**Book Reviews**

**The Supreme Court in Crisis: A Good Read, But No Crisis**

*by Donald A. Daugherty, Jr.*

A Review of:
American Justice 2017: The Supreme Court in Crisis, by Kimberly Robinson

Note from the Editor:
This book review discusses and critiques Kimberly Robinson’s book reflecting on the 2016 Term of the Supreme Court.

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 position, as here, we offer links to other perspectives on the issue, including ones opposed to the position taken in the article. To join the debate, please email us at info@fedsoc.org.


---

**About the Author:**

Don Daugherty is a shareholder in the Litigation Practice Group in the Milwaukee office of Godfrey & Kahn S.C.
In such decisions, the Court often warns in dicta that without grounds in order to avoid wreaking immediate, widespread havoc; which discusses the Roberts Court's so-called “one last chance” implementation of the revised executive order. In a 7-2 vote last December to stay any judicial order enjoining perfectly acceptable if taken by previous presidents, culminating by inferior courts to executive actions that would have been an early example of the Court pushing back on resistance legality. At the same time, the partial reinstatement may have fide relationships” with American citizens or entities was likely a exception created by the Court for individuals having “bona Muslim countries. Robinson writes that the vagueness of the into the United States of individuals from six predominantly Trump's executive order temporarily barring entry begins with the partial reinstatement on the term's very last day are reviewed in “The Priceless Value of Citizenship.” The chapter otherwise have had. However, this does not mean the Court itself race, and drew support for candidate Trump that he might not view of his unyielding opposition. It is generally acknowledged everything about our country has been “in crisis,” at least in the 2016 T erm seems not to have arisen until after the Supreme Court must address. Only a single circuit had considered the matter at the time certiorari was sought in Gloucester County, and no circuit split has arisen since; even with nine justices, the Court presumably will continue to let these issues percolate in the lower courts for the foreseeable future. That the book leads with this example supports the conclusion that there was no real crisis in the 2016 Term.

The notion that the Court was in crisis through the end of its term in June 2017 seems not to have arisen until after the unexpected election of Donald Trump as President in November 2016. Before then, there was little talk of any crisis; since then, everything about our country has been “in crisis,” at least in the view of his unyielding opposition. It is generally acknowledged that the vacancy on the Court focused public attention on the importance of the judiciary as an issue in the 2016 presidential race, and drew support for candidate Trump that he might not otherwise have had. However, this does not mean the Court itself was experiencing any crisis in the 2016 Term.

The 2016 Term’s large number of immigration-related cases are reviewed in “The Priceless Value of Citizenship.” The chapter begins with the partial reinstatement on the term’s very last day of President Trump’s executive order temporarily barring entry into the United States of individuals from six predominantly Muslim countries. Robinson writes that the vagueness of the exception created by the Court for individuals having “bona fide relationships” with American citizens or entities was likely a product of sharp disagreement among the justices about the order’s legality. At the same time, the partial reinstatement may have been an early example of the Court pushing back on resistance by inferior courts to executive actions that would have been perfectly acceptable if taken by previous presidents, culminating in a 7-2 vote last December to stay any judicial order enjoining implementation of the revised executive order.

The book’s most interesting chapter is “Courting Politics,” which discusses the Roberts Court’s so-called “one last chance” doctrine for resolving tough constitutional issues on narrow grounds in order to avoid wreaking immediate, widespread havoc; in such decisions, the Court often warns in dicta that without some legislative or other non-judicial fix, the outcome could be different the next time it is faced with the issue. Robinson argues that decisions causing momentous, social disruption risk exposing the Court to criticism that it is merely another political actor. Beginning with Justice Owen Roberts “switch in time” that mooted FDR’s court-packing plan and ended the Lochner era, Robinson cites other possible examples of a politicized Court—Bush v. Gore, National Federation of Independent Business v. Sebelius, Obergefell v. Hodges, and Citizens United v. Federal Elections Commission. At the same time, although the Court is sometimes fairly accused of deciding issues better left to the political branches or the states, this was not a problem in the 2016 Term, and this chapter has little to do with the term specifically. Further emphasizing the anti-climactic nature of the term, Robinson notes at the chapter’s end that, early in 2016, the Court teed up the issue of whether state laws allowing unions to charge nonmembers “agency fees” or to require that they affirmatively object to contributing to support political causes should be struck down under the First Amendment. After Justice Scalia’s death, however, the Court could only reach a 4-4 non-decision in the Friedrichs case, and it is now revisiting the same issue with nine justices in this term’s Janus v. American Federation of State, County, and Municipal Employees, Council 31.

The term’s most significant, lasting development was undoubtedly the confirmation of Neil Gorsuch, and Robinson gives it due attention in a chapter (regrettably) entitled, “The Stolen Seat.” Like the purported sense of a “crisis,” the notion that the seat was “stolen” only arose after Donald Trump pulled off one of the biggest political upsets in American history. Between the time it became apparent that he would be the Republican nominee until early in the morning after Election Day, the only issue in the minds of Democrats regarding filling the ninth seat was whether President Hillary Clinton would renominate Judge Merrick Garland, or nominate someone to his left.

Furthermore, by replacing Justice Scalia with Justice Gorsuch, the GOP merely held serve. An interesting angle that Robinson leaves unexamined is the decision by Senate Democrats to expend political capital (and allow Republicans to justify extending suspension of the judicial filibuster to the Supreme Court) resisting the exchange of one conservative justice for another. Arguably, Democrats should have kept their powder dry for greater credibility if and when a moderate or liberal justice needs to be replaced.

Rather than an unprecedented “political hijacking” of the Court, as Robinson calls it, Senator McConnell’s refusal to even consider Judge Garland’s nomination simply made things easier for his GOP colleagues, as he drew to himself all the accompanying criticism. It is unlikely that a Republican Senate would have confirmed Judge Garland in the twilight of President Obama’s tenure (unless it was to deny the seat to a more liberal justice whom President Clinton might nominate). This is consistent with the conventional wisdom that a justice must resign early in a President’s term if she wants him to be able to pick her successor; as Justice Ruth Bader Ginsberg explained while responding to calls for her to retire in Fall 2014, President
Obama at that point “could not successfully appoint anyone I would like to see on the court.”

The book closes with a discussion of Justice Gorsuch’s first months on the Court, and then previews the 2017 Term. Robinson observes that early on, Justice Gorsuch showed “he would not shy away from acting on his opinions or fail to make his preferences known;” to her credit, however, she avoids caricaturing Justice Gorsuch as a version of the ambitious, overeager Tracy Flick from the movie *Election*, as some legal commentators have tried to do. Rather, Robinson writes that his genteel yet folksy style during his confirmation hearing was endearing (particularly in contrast to the bitter, partisan approach of many Senators), and that given his academic credentials and judicial experience, “it was hard to poke holes” in his qualifications.

Even the final chapter’s title—“The Calm Before the Storm”—is at odds with Robinson’s contention that the Court was in crisis during the 2016 Term. Looking ahead, Robinson describes the many high profile cases currently before the Court. Besides *Janus* and challenges to the third version of the temporary travel ban, the Court is now considering important post-*Obergefell* issues arising under the First Amendment in *Masterpiece Cakeshop, Ltd v. Colorado Civil Rights Commission*, as well as the political blockbuster *Gill v. Whitford*, which could curb partisan gerrymandering and drastically change how states approach redistricting. Although there is no dispute that the 2017 Term will exceed its predecessor in excitement and controversy, however, Robinson succeeds in turning a sleepy term into an interesting read, even without any real crisis.
Privatization: Boon to Efficiency or Slow Motion Revolution?

by Ted Hirt

A Review of:
Constitutional Coup: Privatization’s Threat to the American Republic, by Jon D. Michaels

http://www.hup.harvard.edu/catalog.php?isbn=9780674737730

Note from the Editor:

The reviewer discusses a new book’s critique of various forms of privatization in the federal government, then presents his own critiques of some of the book’s premises and proposed solutions.

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.


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.

Should our federal government be operated like a business? In Constitutional Coup: Privatization’s Threat to the American Republic, Jon D. Michaels, Professor of Law at the UCLA School of Law, contends that citizen demands that the federal government “be run like a business” are not only ill-founded, but dangerous to our democracy. He also sounds the alarm about the government contracting out key government functions to unaccountable private actors. Ultimately, however, his principal proposed solutions to the problems he identifies simply offer more of the same: an expansive and too often ineffective and inefficient federal government.

Constitutional Coup is a detailed critique of “the almost evangelical denunciation” of the twentieth century administrative state, which Professor Michaels describes as a “multigenerational campaign to refashion public government in the image of a Fortune 500 company, if not now something straight out of the new gig economy.” Michaels defines privatization as “government reliance on private actors to carry out State responsibilities; government utilization of private tools or pathways to carry out State responsibilities; or government ‘marketization’ of the bureaucracy, converting civil servants into effectively privatized, commercialized versions of their former selves and relying on them to carry out State responsibilities.”

Michaels is addressing a substantial issue. In 2009, for instance, the federal government contracted for over $500 billion in goods and services. But this field is already subject to some regulation. For example, in September 2011, the Office of Federal Procurement Policy of the Office of Management and Budget provided a description of “inherently governmental functions” that should not be outsourced, specifically those which “involve the exercise of sovereign powers of the United States,” or, in some cases, that commit the government to a course of action. Clearly this book addresses an important topic, but not one that has been ignored; Michaels adds his voice to a chorus of critiques and proposed solutions.

I. The History of Privatization

Professor Michaels begins by describing our country’s historical experience with the privatization enterprise. In the early republic, the federal government relied on private actors to perform governmental functions, in part out of necessity, and also because the American people had a cultural legacy of distrust of government as a result of colonial experiences under the British
The new country relied on privateers to augment its naval forces, and private bounty hunters supplemented sometimes weak public law enforcement. Private contractors—ranging from the Pinkerton detective agency to private company-operated police forces—aided the public sector in providing security and conducting criminal investigations. The Postal Service relied extensively on private carriers. In an era of limited government functions, with no public welfare or worker protection systems in place, the citizenry did not demand much from its government. Furthermore, that government was staffed by party loyalists, not professionals or technocrats.

Over time—particularly during the early and mid-twentieth century—the federal government expanded the scope of its powers and increasingly relied on government employees rather than contractors to perform government functions. At the same time, this growing government workforce became more professionalized and nonpartisan. These changes enabled the modern administrative state to provide a wide range of social services and support a large military. Congress in turn delegated much of its lawmaking power to the growing administrative agencies, and some private sector systems like industry self-regulation and community charities were displaced.

Michaels defends the constitutionality of the modern administrative state against modern detractors by looking back to the nation’s founding. He explains that the “constitutional separation of powers was deliberated and generally celebrated,” and designed to preclude a concentration of power. James Madison prescribed a “rivalrous separation in which ambition would counter ambition.” Throughout the eighteenth and nineteenth centuries, there was a very small administrative state, with limited powers. But in the 1930’s, the New Deal empowered an array of agencies to expand their powers into the economy, with the Supreme Court’s eventual acquiescence. With the advent of these new agencies, some of the functions that were previously contracted out to private parties were conducted in-house.

The modern administrative state’s power “has long been divided” among “three sets of rivals”—presidentially appointed agency heads, “politically insulated” civil servants, and the general public insofar as it is empowered to participate in the development and implementation of agency policies. Within that structure, the agency leaders try to achieve presidential policies, even at the expense of “rational, legalistic, inclusive, or procedurally robust public administration.” In contrast, the career civil servants “are legally, culturally, and practically independent” like federal judges, and they can resist those policies. Members of the general public, through public comments under the Administrative Procedure Act, lawsuits under the Freedom of Information Act, and other kinds of litigation, can be “constructive participants” in the operations of the administrative state; these efforts are analogous to their separate influence on Congress, although without the corresponding power to recall officials through elections.

Michaels’ thesis is that disagreement or friction among these three factions creates an “administrative separation of powers” that “roughly reproduce[s]” the constitutional rivalries among the legislative, executive, and judicial branches. Michaels contends that these internal conflicts legitimize the administrative state because they restrain the “initially unfettered administrative juggernaut.”

II. Privatization as a Threat to the Administrative State

Michaels recognizes a “growing disenchantment” with the “pax administrativa” among libertarian scholars who criticize the bureaucracy as inefficient, business interests that prefer the rationality of markets, and even elements of the New Deal coalition who have abandoned some of their faith in government. He also detects a confluence of events that facilitated the development of privatization initiatives, ranging from the anti-government rhetoric of Reagan presidency to the distrust of government that is a legacy of the Watergate scandal and the Vietnam War.

Privatization also “grew into a mainstream, bipartisan, and avowedly nonideological movement,” a trend that was promoted by some scholars and ultimately by the Clinton-Gore White House, which touted privatization as a “technocratic endeavor.” Privatization extends beyond the contracting out of government services; it includes private standard setting, private accreditation and administration, the marketization of the bureaucracy (e.g. transforming career civil servants into at-will employees).
employees), the deputization of private actors for government intelligence or law enforcement, and equity investing and “back door” regulation (e.g. the bailouts that occurred in the aftermath of the 2008 financial crisis). 31 Michaels is particularly concerned about the contracting out of policy administration and the military’s reliance on contractors. 32 Other commentators have likewise questioned the military’s use of contractors in conjunction with the operation of the Abu Ghraib prison in Iraq and in the Iraq War more generally. 33

Michaels’ objection to privatization transcends the ongoing debate over whether privatization is more economically efficient than relying on government personnel to deliver services, or whether privatization ultimately produces a leaner government. 34 Instead, Michaels contends that privatization permits the government to expand its powers “at the expense of the private sector,” destabilizing the “liberal democratic order.” 35 How might that happen? Michaels argues that, where political leaders can accomplish their goals through contractors without the input or resistance of the civil service, the check that would otherwise be provided by civil servants in the administrative separation of powers is bypassed, and the balance provided by that scheme is thereby tilted in favor of political appointees. 36 Contractors are also more likely to disregard procedural or other constraints on their conduct that would apply to government employees carrying out the same tasks, and they are financially motivated to carry out their responsibilities more quickly (for good or ill). 37 In addition, contracting “marginalizes public participation in the administrative process” because the public cannot effectively monitor contractors in the way it can monitor agency employees. 38

Michaels rejects the notion that “gridlock” in government is necessarily bad, acknowledging that the nation’s inability to resolve difficult issues may reflect the absence of a consensus or our culture’s “cautious approach to governing and . . . reluctance to regulate or legislate unless and until broad-based buy-in is secured.” 39 Efficient or not, Michaels views the administrative separation of powers he posits as a bulwark against authoritarianism. 40

III. Pushing Back Against Privatization

Michaels wants the various power centers within the administrative state to remain dispersed. He opposes delegations of congressional authority directly to the President, which he describes as a “naked delegation” of power. 41 In contrast, Michaels characterizes delegations to agencies as less problematic insofar as the agencies contain their own sources of power and disagreement and, in turn, must work out differences with members of civil society. 42

Judicial review under the Administrative Procedure Act (APA) permits federal judges to invalidate agency action that is, among other things, “contrary to constitutional right, power, privilege, or immunity,” that violates a statute or statutory procedures, or that is “arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.” 43 In contrast to the current dual review of both procedural and substantive (“arbitrary and capricious”) decisions, Michaels wants judicial review to focus on the processes by which rules and other agency decisions are reached. 44 Michaels believes that courts can support the administrative separation of powers by scrutinizing the extent to which agency actions were reached “through a truly rivalrous, heterogeneous, and inclusive administrative process.” 45 He reasons that courts “are more adept at identifying shoddy procedures than they are at assessing suspect substantive policy decisions, particularly in cases involving heavy reliance on sophisticated scientific, sociological, or economic analysis.” 46 Michaels basically rejects the need for judicial review of the merits of agency decisions if the process is legitimate and if no statute has been violated. 47

But Michaels urges more expansive judicial review for situations in which, according to his view, the “spirit” of the administrative separation of powers has been violated, even where there has not been a violation of specific statutes or judicial orders. 48 For him, this spirit is violated whenever an agency makes a move toward privatization. He does not spell it out, but perhaps he would want the APA amended to empower the courts to review and reject, for example, the outsourcing of rule-drafting, decisions on whether individuals are eligible for government benefits, or granting interest groups special roles in the formulation of government policies. 49

Finally, Michaels identifies a separate category of what he considers to be “hard cases” to which judicial review could be

---

31 Id. at 106-11.
32 Id. at 112-14.
34 Michaels, supra note 1, at 120-23.
35 Id. at 126.
36 Id. at 128-29.
37 Id. at 132.
38 Id. at 131.
39 Id. at 147-48.
40 Id. at 150.
41 Id. at 176.
42 Id.
44 Michaels, supra note 1, at 177.
45 Id. at 181. In that context, Michaels cites United States v. Mead Corp, 533 U.S. 218, 229 (2001), in which the Court withheld so-called Chevron deference (articulated in Chevron, U.S.A., Inc. v. Natural Resources Defense Council, Inc., 467 U.S. 837, 841-43 (1984)), from an agency decision, according to Michaels, “on the ground that career agency personnel formulated their own legal interpretations outside of the democratically inclusive and rigorous rulemaking process—and without apparent input from agency leaders or the public writ large.” Michaels, supra note 1, at 183.
46 Id. at 185.
47 Id. at 185, 180.
48 Id. at 189.
49 Id. at 189-90.
more difficult to apply. These hard cases include federal efforts to marketize the bureaucracy, to create independent agencies, and to exempt national security agencies and government corporations from many of administrative law’s routine procedural and personnel safeguards.50 Michaels indicates that, despite the difficulty, judicial review should apply to all of the activities in this category except national security agency activities.51

Michaels advocates other reforms to reverse privatization.52 He proposes a “comprehensive civil service reclamation project” that would focus on “renationalization, reinstatement, recruitment, retention, and reputation building.”53 He wants Congress and the President to reverse outsourcing, at least where contractors or other entities have a substantial role in making or implementing policy.54 Michaels also urges the creation of a national government service academy analogous to the military’s professional academies and public campaigns to support such career choices.55

Finally, Michaels wants to see civil society strengthened so that it can have an effective stake and voice in the administrative state.56 He would supplement the “public notice” of government activities prescribed in the Federal Register through both mainstream media and internet links, with enhanced use of “plain English” and better support for “lay comprehension” of agency documents, and increased agency use of social media.57 Agency officials also would engage in more frequent community outreach and expanded civics programs.58

IV. Critique

Although Michaels’ defense of the administrative state and critique of privatization is both articulate and well-argued, I have several concerns about both the accuracy of his premises and the wisdom and practicality of his proposed solutions.

Michaels proceeds from the premise that an administrative separation of powers exists under which the civil service bureaucracy, presidential power, and public participation check each other.59 But Michaels does not provide enough evidence to support the notion that there is an inherent balance of power among the three factions, and there are reasons to think there is not. The interests of the civil service bureaucracy are constant, while the interests of the political leadership of agencies change depending on who is elected president; the interests of the two factions are sometimes opposed—but they are sometimes aligned. Recent data indicates that federal civil servants overwhelmingly vote for and donate to Democrats,60 and they tend to be interested in maintaining their own jobs and expanding the scope of their influence. While an administration seeking to shrink government might be effectively checked by reluctant civil servants, an administration committed to maintaining or expanding social welfare or regulatory programs will likely find supporters within the bureaucracy, whose ideological and personal interests support those objectives. Where the interests of these two factions are aligned, there is no reason to conclude that bureaucratic views will “check” administration initiatives.

Second, Michaels’ cry of alarm about the extent of harmful privatization is not sufficiently substantiated. Michaels has not demonstrated that private entities have destabilized our government structure by systematically undermining public policy or frequently acting without public agency oversight.61 There is no doubt that some abuses of authority have occurred; the military contractors’ operation of the Abu Ghraib prison is a prominent example.62 But Michaels has not established that privatization initiatives have substantially undermined governmental accountability. In order to do so, Michaels would have to show that the private actors have systematically engaged in activities they were not legally empowered to conduct (either by law or by the terms of their contracts), or that they regularly pursued lawful goal in unlawful ways.

Third, it is not clear that Michaels’ proposed reforms to judicial review of agency decisions would ameliorate the privatization problem. As explained above, Michaels eschews broad judicial review of the merits of agency decisions except where they implicate private entities. Where private entities come into play, Michaels would expand judicial review beyond an evaluation of the legality of the agency’s ultimate decision, which is ordinarily where judicial review begins and ends today.63 Enhanced scrutiny of contracting decisions presumably would extend deeply into agencies’ internal decision making processes. Such intrusion would be disruptive (if not demoralizing) to the functioning of the administrative state that he supports.64 We also should be concerned about empowering judges to evaluate the legitimacy of outsourcing initiatives, for it is not clear what, if any, objective standards would guide their review of agency decisions.

---

50 Id. at 193-97.
51 Id. at 193-200.
52 Id. at 202-18.
53 Id. at 206.
54 Id. at 207.
55 Id. at 209-18.
56 Id. at 218.
57 Id. at 220-22, 224-26.
58 Id. at 222, 227.
59 Id. at 59-68.

61 Michaels, supra note 1, at 126.
62 See Etzioni, supra note 33, at 310-11.
64 In that context, “exemption 5” of the Freedom of Information Act, 5 U.S.C. § 552(b)(5), protects from public disclosure the internal deliberations of government officials. See New Hampshire Right to Life v. U.S. Dep’t of Health & Human Services, 778 F.3d 43, 52 (1st Cir. 2015) (Exemption 5 protects government agencies from being “forced to operate in a fishbowl.”) (citations and internal quotation marks omitted).
Aside from substantive concerns about these proposals, it is not realistic to think that Congress would rewrite the APA to curtail arbitrary and capricious review of agency action, or substitute for it a broad review of agency decisions related to contracting. 65

A final practical objection to Michaels’ project: government employees currently do not have the expertise needed to engage in standard-setting, information technology services, or other similarly specialized tasks in our complex, technology-based society, so outsourcing such complex tasks is a practical necessity. 66 Absent a commitment by Congress to insource such complex functions through sweeping changes to the federal civil service appointment and compensation system, these functions will necessarily remain the province of private contractors.

Michaels’ concern about whether “inherently governmental functions” should be delegated to private contractors, however, is a legitimate one. 67 Functions associated with sovereignty—such as national defense and intelligence-gathering—and policymaking must be kept in-house. Conservatives who are concerned about the unaccountable exercise of government powers should be sympathetic to increased scrutiny of such practices. 68 And to the extent the services of outside contractors are not sufficiently monitored, improvements should be made, including by expanding the number of contracting officers or auditors if necessary. 69 Michaels’ proposed improvements to civil service functions and his ideas for enhancing public participation may also be worth trying. Michaels identifies some real concerns and proposes some promising policy shifts, but his larger vision is not a satisfactory response to the problems he identifies.

---

66 See Emily S. Bremer, Private Complements to Public Governance, 81 Mo. L. Rev. 1115, 1121-22 (2016) (estimating that there may be 100,00 private technical standards in place); United States Government Accountability Office, Information Technology, Agencies Need to Involve Chief Information Officers in Reviewing Billions of Dollars in Acquisitions (January 2018) at 2 (reviewing aspects of IT contracts of over $19 billion for fiscal year 2016).
68 See Department of Transportation v. Ass’n of American Railroads, ___ U.S. ___, 135 S. Ct. 1225, 1234 (2015) (Alito, J., concurring) (“When citizens cannot readily identify the source of legislation or regulation that affects their lives, Government officials can wield power without owning up to the consequences. One way the Government can regulate without accountability is by passing off a Government operation as an independent private concern.”).
The Federalist Society Review
Volume 19

REGULATING UNDER THE RULE OF LAW

by J. Kennerly Davis

A Review of:
How to Regulate: A Guide for Policymakers,
by Thomas A. Lambert

Note from the Editor:
The reviewer discusses a book about regulation, critiques it for perpetuating the wrongheaded assumptions that underlie our existing regulatory system, then proposes a path forward to bring our regulatory system into conformity with the rule of law.

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.


About the Author:

J. Kennerly Davis is a former Deputy Attorney General for Virginia and a contributor to the Federalist Society’s Regulatory Transparency Project.

Government regulation defines our daily existence. Regulation defines our education, employment, retirement, health care, and much more. Using comprehensive systems of technical specification, procedural requirements, permitting, and licensing, with close monitoring and stringent penalties for noncompliance, regulation determines the distinguishing characteristics of the products we buy and the essential terms and conditions of the services we purchase. Regulation even seeks to govern what we say, what we write, and the circumstances under which we may express our religious beliefs.

All of this is done in the pursuit of policy objectives that are widely supported, at least as those objectives have been stated generally: clean air and water, a protected environment, a safe work place, affordable housing and health care, diversity, and tolerance. The overall system of regulation that has been established to pursue these and other generally laudable objectives is so extensive and so complex that no one can with certainty identify all the agencies, departments, commissions, bureaus, and boards that manage our affairs and direct our activities.

The regulatory system is not only pervasive and complex, it is enormously expensive. Some analysts estimate that the cost to comply with government regulation now totals approximately $2 trillion per year.1 To cover their compliance costs, American businesses are forced to raise the prices they charge for the goods and services they sell. Businesses are also forced to cut their other costs, reducing what they can spend on wages, employee benefits, research and development, and production facility upgrades.

In this way, government regulation imposes an enormous burden on the American economy, a hidden tax that we all must pay in higher prices, smaller paychecks, reduced benefits, and lost opportunities for employment. If this hidden tax were spread evenly to all American households, the share for each household would amount to approximately $15,000 per year—an amount fifty percent larger than all the other taxes the average household has to pay in the year!2

Supporters of extensive regulation dispute the $2 trillion figure, arguing that compliance costs are much smaller, or in any event smaller than the value of the significant social welfare benefits that are provided by regulation. Disputes about the costs and benefits of regulation are likely to persist because there is no general agreement about how to calculate those costs and benefits.

Given the pervasiveness, complexity, and cost of government regulation, Americans have a vital interest and real need to understand as best they can how the regulatory system currently works and how regulation can be made to achieve its laudable objectives effectively in a way that minimizes compliance costs and undesirable side effects.

I. The Author’s Approach

Thomas Lambert seeks to meet the pressing need we have to understand the regulatory system, in his 2017 book, How to

---

1 See Clyde Wayne Crews, Jr., Ten Thousand Commandments 2018 (Washington, D.C.; Competitive Enterprise Institute, April 2018).
A. What is Regulation?

At the outset of this book, Professor Lambert observes that the general concept of regulation—control and direction of human behavior—is quite broad. It includes a wide variety of common law doctrines such as nuisance, trespass, and negligence, as well as privately agreed to arrangements ordered by doctrines of contract law.

But Professor Lambert explicitly excludes from his concept of regulation all the common law doctrines of property, tort, and contract that provide for what he calls “private ordering.” Indeed, he thinks of regulation in traditional progressive terms as a system of governmental control and direction based on statutes and agency rules designed to correct the defects in private ordering—so called “market failures.” As he puts it:

A workable definition of regulation would be any threat-backed governmental directive aimed at fixing a defect in “private ordering”—the world that would exist if people did their own thing without government intervention beyond enforcing common law rights to person, property and contract—where the defect [the market failure] causes total social welfare (i.e., the aggregate welfare of all citizens) to be lower than it otherwise would be.5

If regulation is defined this way, as a supplement and correction to private ordering arrangements based on common law, then Professor Lambert offers that regulation will always be dealing with what he calls “mixed-bag” conduct—human behavior that sometimes and to some extent is good, and that sometimes and to some extent is bad. By “good” he means behavior that increases total social welfare, and by “bad” he means behavior that decreases total social welfare. Regulation must deal with mixed-bag behavior because common law doctrines have long condemned behavior that is always bad (e.g., battery or fraud) and regulation must take over where the common law leaves off.6

If the purpose of regulation is to govern mixed bag behavior to increase total social welfare, then Professor Lambert proposes that the goal of policymakers should be to “craft legal directives so as to prevent the bad aspects of mixed-bag behavior without simultaneously forbidding or discouraging the good aspects[,]”7

This is not a simple task. There is always the danger that the decisions of policymakers will err in one of two directions. They may fail to prohibit some kinds of bad conduct or they may prohibit some kinds of good conduct. Either type of error creates what Professor Lambert calls “error cost” that results in a lessening of total social welfare. To minimize the possibility of error cost, policymakers can make their regulatory directives more “nuanced.” Policymakers accomplish this by spelling out their directives in great detail using text and terminology that relies upon slight and finely shaded distinctions to specify the precise circumstances in which a behavior will be regulated and, if regulated, the precise circumstances in which the behavior will be permitted or prohibited.

Making regulatory directive more nuanced may help reduce error cost, but Professor Lambert admits that adding nuance will increase what he calls the “decision costs” associated with the directives. In the first place, it’s more costly for regulators to formulate directives that are very detailed and complex. It’s also more costly for regulated parties to interpret and comply with directives that are very detailed and complex. And it’s more costly for enforcement officials and adjudicators to apply such directives and determine in a particular case whether the behavior in question is permitted or prohibited.

There are difficult trade-offs posed by the intertwined problems of error cost and decision cost. Any steps that policymakers may take to reduce one kind of cost may increase the other kind of cost. If policymakers reduce the coverage of a directive to lessen the risk of prohibiting good behavior, then bad behavior may increase. If, on the other hand, policymakers increase the coverage of a directive to lessen the risk of permitting bad behavior, then good behavior may decrease. And if policymakers

---


4 Id. at x, 2.

5 Id. at 4.

6 Id. at 7, 8.

7 Id. at 8.
attempt to reduce error costs by issuing directives that are more nuanced, then decision costs will certainly increase.

In light of the trade-offs and tension that result from the problems of error cost and decision cost, Professor Lambert proposes a rule of thumb for policymakers. He suggests that they should always strive to "craft directives to minimize the sum of error and decision costs" so as to maximize the net social welfare benefits produced by the directives. 8

B. The Physician Model

Professor Lambert recognizes that his rule of thumb may, at first glance, appear to call for nothing more than very basic cost-benefit analysis whereby policymakers simply estimate the compliance costs and social welfare benefits of a particular proposed regulation and then issue the directive if its estimated benefits exceed its estimated costs.

But he asserts that his rule of thumb actually calls for a more sophisticated form of cost-benefit analysis that, in addition to considering the costs and benefits directly resulting from a particular regulatory proposal, also considers that costs and benefits that would result from the implementation of regulatory alternatives. In this way, policymakers can ensure that they are "taking account of the opportunity costs of selecting one regulatory approach over its alternatives." 9 Policymakers will then "have some sense of how net welfare would differ under all regulatory alternatives." 10

Professor Lambert recognizes that the kind of comprehensive analysis he calls for can easily become quite complex and quite costly. To keep things reasonably in check, he advises policymakers "to gather [only] as much welfare-related information as can be cost-effectively gathered on all potential regulatory options, then select the option that minimizes the sum of error and decision costs." 11

It’s fine, of course, to say that policymakers should perform their comprehensive analysis of regulatory alternatives cost-effectively, but how exactly should they proceed to do so? Professor Lambert suggests that policymakers crafting regulations should follow the lead of physicians who, he argues, routinely perform the type of analysis that he is proposing for policymakers. A doctor seeing a patient with an ailment aims to select not just a remedy that will improve the patient’s health (i.e., a remedy that will create more benefits than costs), but the remedy that will maximize net benefits for the patient. The doctor needs to consider the treatment alternatives and select one quickly and efficiently.

According to Professor Lambert, the doctor performs this analysis using a five-step process. First, the doctor identifies the patient’s symptoms—the adverse condition that varies from the healthy norm. Next, the doctor diagnoses the disease that has caused the symptoms—the source of the adverse condition. Then, the doctor identifies the alternative remedies available to cure the illness or at least alleviate the symptoms. Then, the doctor assesses all the costs and benefits associated with each available remedy—things like monetary cost, potential side effects, and the expected effectiveness of each remedy. Finally, the doctor selects the remedy that will provide the greatest net benefits to the patient.

"The central claim of this book is that policymakers . . . should follow the lead of physicians." 12 This is the essence of Professor Lambert’s “unified practical approach to regulating.” 13 Policymakers should craft their regulatory directives using the same five-step process physicians use to formulate treatment plans for sick patients.

First, policymakers must identify the symptom—“the adverse effect citizens confront within the scheme of private ordering,” the reduced total social welfare that exists as a result of people “doing their own thing” in accordance with their private ordering arrangements based on common law doctrines of property, tort, and contract, and without the benefit of government regulatory intervention. 14

Such symptoms will be prevalent and persistent because, according to Professor Lambert, private ordering arrangements based on common law doctrines can never maximize total social welfare to the extent policymakers can through the implementation of their regulatory directives. Professor Lambert concedes that common law private ordering theoretically maximizes total social welfare. That is because rules protecting people’s bodies and property and allowing them to enforce contracts provide strong support for voluntary association, commercial exchange, and economic development:

[W]hen property rights are well defined and transferrable, and individuals are able to strike trustworthy exchange agreements, markets will emerge and channel productive resources to their highest and best ends—production of the goods and services that individuals value the most. 15

Unfortunately, as Professor Lambert sees it, “markets sometimes fail to work well” when they are regulated only by the common law. 16 Indeed, he sees markets failing on a regular basis because “people left to their own devices, constrained only by the common law . . . systematically fail to extract the greatest possible value from available resources.” 17 As a result of these systematic market failures, policymakers are confronted with a great many symptoms of suboptimal social welfare.

Once they identify a symptom, policymakers can proceed to work through steps 2–5. They must seek to diagnose the cause of the symptom to identify the reason that private ordering arrangements and market mechanisms have failed to maximize total social welfare. Once they have diagnosed the market failure causing the symptom, policymakers then must cost-effectively catalogue the different regulatory remedies available to treat

8 Id. at 12, 13.
9 Id. at 13.
10 Id. at 14.
11 Id.
the disease that they have diagnosed. Some remedies may only alleviate the symptom; others may address the market failure itself. After cost-effectively cataloging the available regulatory remedies, policymakers must finally “make an informed judgment as to the net benefits of each” and select the regulatory alternative that will “minimize the sum of decision and error costs.”

Having set forth his physician’s approach to regulating in general terms, Professor Lambert devotes the rest of his book to detailed descriptions of how that approach can be applied to six different defects in private ordering arrangements—six diseases that are often used to justify the imposition of regulatory directives. These are the often discussed market failures associated with (i) unwanted economic externalities, (ii) public goods that are overconsumed and under-produced, (iii) excessive market power, (iv) information asymmetry among market participants, (v) agency costs that result from excessive market power and information asymmetry, and (vi) cognitive limitations and behavioral quirks that prevent market participants from extracting the greatest possible value from available resources. For each of these defects, Professor Lambert discusses the symptoms they present, the sources of those symptoms, the remedies that are available, and how policymakers can apply his physician’s approach to select the best regulatory remedy—the remedy that seems likely to minimize the sum of error and decision costs and thereby to maximize total social welfare.

II. Assessing the Author’s Approach

Professor Lambert’s stated goal is to systematize the insights of a variety of legal theorists and economists in a single non-technical book that offers a unified practical approach to regulating: the physician’s approach. He aims to educate policymakers, and especially to broaden the professional perspectives of lawyers and economists directly involved in the regulatory process, so they can make “better regulatory decisions that produce greater human welfare” and “improve regulatory performance.”

In short, Professor Lambert aims to make regulation better. Does he succeed? Given the pervasiveness, complexity, and cost of the current regulatory system, it is impossible not to support the efforts of Professor Lambert, or anyone who wants to make regulation better. But given the importance of the subject, we must give very careful consideration to the question of whether, how, and to what extent his physician’s approach would actually improve the regulatory system.

“Better” is a relative concept that calls for the comparative evaluation of the attributes of two things in light of some standard that defines the attributes of the optimal version of the two things being compared. So we must first consider the standard Professor Lambert sets forth for regulation. Then we can evaluate whether his approach makes regulation better.

As we have seen, Professor Lambert defines regulation as any threat-backed top down governmental directive imposed by policymakers on private ordering arrangements to correct what the policymakers see as defects in those arrangements, so-called market failures that the policymakers believe have caused total social welfare to be lower than it would be absent the market failures. Typically, such directives come from executive branch policymakers who are presumed to be experts regarding the subject matter they regulate.

For Professor Lambert, the optimal regulatory system is one in which the policymakers have all the information they need to thoroughly analyze the market failure in question, and also to thoroughly evaluate and compare all the alternative regulatory responses that might be imposed to correct the failure. With this comprehensive information, the policymakers can apply their technical expertise and professional judgment to select and impose the regulatory correction that will maximize total social welfare while minimizing the error costs and decision costs that will result from the imposition of the regulation. The optimal regulatory system, therefore, is one that enables all-knowing policymakers to maximize benefits and minimize costs. With this standard in mind, how does Professor Lambert’s physician’s approach to regulation compare to the possible alternatives?

A. Just the Status Quo?

In the concluding chapter of his book, Professor Lambert asserts that “[a]dherence to the [physician’s approach] would certainly represent an improvement over the regulatory status quo, in which policymakers regularly leap from identification of a symptom [of market failure] to implementation of a remedy.”

He cites and criticizes the net neutrality rules of the Obama administration as a case where “policymakers bypassed altogether the sort of analysis this book recommends” and “just opted for a politically expedient power grab.”

Professor Lambert’s point, while true enough, is rather limited in its significance. His physician’s approach is certainly better than an approach to regulating based solely on the desire of policymakers to amass political power. That said, it is far from clear that policymakers regularly impose new regulations without any analysis.

In fact, a longstanding series of Executive Orders require executive branch agencies to engage in pre-regulation analysis very similar to that embodied in Professor Lambert’s physician’s approach to regulating. For example, these Executive Orders require agencies to identify the market failures that they intend to address, identify and assess alternative forms of regulation, assess all the costs and benefits of available regulatory alternatives, select the approach that maximizes net benefits and, throughout the process, design regulation in the most cost-effective manner possible.

At one point, Professor Lambert asserts that his physician’s approach “improves upon the simplistic cost-benefit analysis often invoked in debates over proposed regulations.” Considering the forests that have been felled to supply the paper needed to record the endless debates about the regulatory analysis performed by

18 Id. at 15.
19 Id. at x, 2.
20 Id. at 251.
21 Id.
23 Lambert, supra note 3, at 13.
many, if not all, agencies to comply with the aforementioned Executive Orders, Professor Lambert is closer to the mark when he acknowledges that the regulatory analysis currently performed by agencies already “looks an awful lot like” his physician’s approach.24

Professor Lambert admits that while he was writing he sometimes “wondered whether the book was ‘original’ enough.”25 He confides that he took comfort from the example of C.S. Lewis, “whose *Mere Christianity* systematized the ideas of scores of philosophers and theologians, spoke plainly, and became a classic.”26

Professor Lambert’s systemized exposition of the different types of market failures, and the regulatory alternatives available to respond to those failures, incorporating as it does the ideas of a wide range of legal theorists and economists, certainly provides a useful compendium of traditional thoughts about regulation. Lawyers, especially, may find Professor Lambert’s generally accessible discussion of economic theory to be useful. Only time will tell whether this useful book also becomes a classic. But for now, it seems reasonable to conclude that his physician’s approach to regulating is not clearly better than the status quo.

B. Progressive Assumptions About State and Society

This lack of distinction from the status quo stems from the fact that Professor Lambert’s approach to regulating and the standard he defines for the optimal regulatory system incorporate all the basic concepts and assumptions that have characterized traditional progressive thought and theories about regulation since the emergence of the administrative state more than 100 years ago. Indeed, it is worth noting that even the illustration on the cover of Professor Lambert’s book is faithful to progressive tradition. “The Regulator” is depicted in the style of Social Realism that was common in the 1920s and 1930s when the walls of public buildings from San Francisco to Stalingrad were adorned with dramatic images of the muscular masses toiling away heroically in mines and fields and factories. Here, the Regulator is shown struggling to restrain an exuberant horse, perhaps one that just wants to do its own thing.

At the center of the progressive regulatory system stand the all-knowing expert policymakers— schooled in the social sciences, law, and public administration, fully informed, disinterested, technocratic—operating outside the realm of politics, freed by a reinterpreted Constitution and broadly written statutes to exercise their professional judgment to remedy any market failures they think warrant their attention. By permitting and prohibiting, rewarding and punishing, allocating and re-allocating the resources of society, these experts use their threat-backed top down executive directives to adjust and recast the arrangements that private persons have made in the belief that such top down reordering will increase the welfare of society as a whole.

Three basic concepts support the traditional progressive regulatory system and justify its operation: (i) the idea that market based arrangements among private parties routinely misallocate resources, (ii) the idea that government policymakers are capable of formulating executive directives that can correct private ordering market failures and optimize the allocation of resources, and (iii) the idea that the welfare of society is something that actually exists separate and apart from the individual welfare of each of the members of society. Each one of these concepts is fatally flawed. None can be relied upon to justify the operation of the administrative state, either as it is or as Professor Lambert wants it to be.

1. Market Failure

There are at least two serious problems with the concept of market failure. First, an enormous body of empirical evidence convincingly demonstrates that market based arrangements among private parties are, in fact, the most efficient and cost-effective means by which to achieve the optimal allocation of the limited resources available to the individual members of a society.27 These market based private ordering arrangements are “defective” only if judged against the subjective preferences that progressive policymakers have for certain resource allocations that differ from the ones that result from private ordering arrangements. For example, the availability of low-cost, high-deductible health insurance providing basic coverage is evidence of a failure in the health insurance markets only if the availability of such a product is judged against the subjective preference that progressive policymakers have for mandatory health insurance providing only comprehensive coverage.

The concept called market failure is also flawed because it distorts any analysis of an economic problem that begins with its presumption. The concept artificially and incorrectly anchors the sequential analysis of an economic problem by first assuming that the problem is the result of some defect in private ordering arrangements that exist in some imagined unregulated state of nature, as opposed to the real world where all transactions are influenced to some degree by regulation. With this assumption in place to direct the analysis, it is logical, even inevitable, for the policymaker to conclude that the proper response to the observed problem is to impose new regulations; conversely, this assumption will never yield a solution that involves removing or narrowing the scope of a regulation. The assumption that problems arise exclusively from failures in private ordering arrangements is the force that drives the often commented upon “ratchet effect” attribute of the administrative state; it only grows, never shrinks. Professor Lambert does note that “[i]n many cases, the optimal remedy [for a market failure] will be to do nothing and simply allow space for privately ordered solutions to . . . emerge.”28 Regulatory forbearance is always welcome, but the fact that a policymaker may, on a case-by-case basis, elect not to initiate a process based upon a distorting assumption of market failure does nothing to address, much less correct, the fundamental flaw in the process that results from that distorting assumption.

---

24 Id. at 252.
25 Id. at xi.
26 Id.
28 Lambert, supra note 3, at 250.
As everyone knows, the existing regulatory system is pervasive, all-encompassing, and controlling. It is, therefore, profoundly mistaken to begin the analysis of an economic problem with the assumption that the problem arose as a result of a lack of regulation. For example, traditional progressive analysis blames the 2008 financial crisis on the greed of bankers, builders, and other private parties. Greed is a facet of human nature and always present to some degree in the actions of private parties. Greed certainly played a part in the financial crisis, but to conclude that greed caused the crisis is to ignore the regulatory failures and flawed public policies—including affordable housing policies, lapses in regulatory oversight, rating agency failures, and cheap money policies advanced by the Federal Reserve—that made significant causal contributions to the crisis.31 Considering the pervasiveness of the current regulatory system, it would be much more accurate, and potentially more useful, if policymakers began their analysis of an economic problem by assuming a regulatory failure.

2. The Expert

As they emerged in the early 20th century, the progressives completely rejected the concept of popular sovereignty based on natural rights that is embodied in the Declaration of Independence. Woodrow Wilson, for one, dismissed the Declaration as a political document of some historical interest, of Independence. Woodrow Wilson, for one, dismissed the Declaration as a political document of some historical interest, based on natural rights that is embodied in the Declaration. Woodrow Wilson, for one, dismissed the Declaration as a political document of some historical interest, based on natural rights that is embodied in the Declaration. Wilson argued that the efficient and effective government that modern America so desperately needed.35 For these wise few to fulfill their destiny and reform the government, Wilson argued that it was essential to "discover the simplest arrangements by which responsibility can be unmistakably fixed upon [these] officials" so as to provide them with "large powers and unhampered discretion" to apply their expertise without political interference from the nominally sovereign but unwise populace.36

Our success [in reforming government] is made doubtful by that besetting error of ours, the error of trying to do too much by vote. Self-government does not consist in having a hand in everything . . . [As in housekeeping, the] . . . cook must be trusted with a large discretion as to the management of the fires and ovens . . . without suffering . . . [public opinion] . . . to be meddlesome.37 Thus was born the concept of the expert policymaker, the professionally trained technocrat, operating apart from politics, applying his superior wisdom and keen insights to identify the defects in the arrangements made by lesser Americans, then exercising his large powers and unhampered discretion to impose threat-backed directives on his lessers so as to override and modify their arrangements in ways he believes will benefit society as a whole.

Wilson’s concept of the wise expert is still of central importance to the progressive regulatory system today. Professor Lambert has directed his book primarily to experts. He has offered up his physician’s approach to regulating in hopes of helping

31 Woodrow Wilson, The Study of Administration, 2 Pol. Sci. Q. 197 (June 1887).
32 Id.
33 Id.
34 Id. at 208-9.
35 Id.
36 Id. at 209.
37 Id. at 214-5.
experts function more efficiently. And Professor Lambert’s optimal regulatory system is centered around fully functioning experts.

This progressive trust in experts is misplaced. It is simply false to suppose that government policymakers are capable of formulating executive directives that effectively improve upon private arrangements and optimize the allocation of resources. Friedrich Hayek and other classical liberals have persuasively argued, and everyday experience has repeatedly confirmed, that the information needed to allocate resources efficiently is voluminous and complex and widely dispersed. So much so that government experts acting through top down directives can never hope to match the efficiency of resource allocation made through countless voluntary market based transactions among the private parties who actually possess the information needed to allocate the resources most efficiently.\(^{38}\)

Professor Lambert acknowledges the work of Hayek and the fact that the information relevant to resource allocation is widely dispersed:

A central theme of this book is that Hayek’s knowledge problem—the fact no central planner can possess and process all the information needed to allocate resources so as to unlock their greatest possible value—applies to regulation, which is ultimately a set of centralized decisions about resource allocation.\(^{39}\)

He recognizes that context-specific information is vitally important. But, having conceded these critical points, Professor Lambert fails to follow them to the logical conclusion that private ordering arrangements are best for allocating resources efficiently. Instead, he stops one step short, suggesting that policymakers defer to the regulator most familiar with the regulated party when they need context-specific information for their analysis.\(^{40}\)

Professor Lambert is mistaken. The best information for resource allocation is not to be found in the regional office of the regulator. It resides with the persons who have long been controlled and directed by the progressive regulatory system. These are the ones to whom policymakers should defer.

Policy directed resource allocations are insufficient for another reason. Actual policymakers simply don’t live up to Wilson’s ideal of the disinterested, objective, apolitical, expert technocrat. To the contrary, a vast amount of research related to public choice theory has convincingly demonstrated that the decisions of regulatory agencies are frequently shaped by institutional self-interest and the influence of the entities the agencies regulate.\(^{41}\) Professor Lambert devotes a good deal of attention to the problem of “agency capture” by regulated entities.\(^{42}\) However, he fails to acknowledge that a symbiotic relationship between regulators and regulated is not a bug in the regulatory system, but an inherent feature of a system defined by extensive and continuing government involvement in the allocation of resources.

Finally, in addition to lacking the information and objectivity needed to allocate resources efficiently, the expert also lacks the legitimacy to allocate resources lawfully. Our constitutional system clearly divides the legislative, executive, and judicial functions among three branches. Wilson and the progressives rejected this constitutional system and consolidated all these functions in the powerful and politically unaccountable executive expert. The Founders realized that such a consolidation of power defined the essence of tyranny.\(^{43}\) They established a system of separated powers to secure their natural rights and protect liberty at the expense of efficiency. The progressives reversed these priorities and instead sought efficiency at the expense of liberty. Considering the limited information available to policymakers, and their revealed lack of objectivity, it appears that we have sacrificed liberty to seek efficiency and now have little of either.

3. Social Welfare

The whole point of the progressive regulatory system—the purpose and justification of every one of its imposed directives—is to increase social welfare. Professor Lambert’s ultimate goal for his book is to provide policymakers with a resource that will enable them to make regulatory decisions that produce greater social welfare. There is, however, a fatal flaw in the concept of social welfare that undermines the legitimacy of any effort to increase it using threat-backed top down regulatory directives. The concept is fatally flawed because there is simply no meaningful way to measure it. There is no single generally accepted methodology that anyone can use to determine objectively how and to what extent the welfare of society will be affected by a particular regulatory directive.

Policymakers have long used gross domestic product ("GDP")—the market value of all goods and services produced in a year—as a measure of the welfare of society.\(^{44}\) But GDP has come under sustained attack from a multitude of critics who argue that it provides an exceedingly narrow measurement of certain types of quantifiable economic activity, and that it fails to adequately account for the substantial social benefits associated with many other factors relating to the quality of life. Those critics contend that these quality of life factors should also be considered when assessing social welfare.\(^{45}\) The problem is that the benefits associated with these additional quality of life factors are extremely difficult, if not impossible, to quantify for use in any measurement of social welfare; consider for example, the benefits associated with atmospheric decarbonization, sustainable economic activity, equitable income distribution, and leisure time.


40 *Id.* at 251.


42 Lambert, *supra* note 3, at 33, 80-84, 233.

43 *The Federalist* No. 47, at 249 (James Madison) (George Carey and James McClellan ed., 2001).


45 See Moss, * supra* note 44, and Ramanujam, * supra* note 44.
When economists attempt to measure social welfare, they begin by selecting the social welfare function or mathematical model they will use for their analysis. Several different models are currently in use; no single model is used universally or exclusively. When the economists have selected their model, they will then select the inputs to be modeled. Inputs may include any factors considered by the economists to have a material effect on the welfare of society. Some of these factors, like GDP, may already be quantified. Other quality of life factors like atmospheric decarbonization are not. To input these factors into their model, economists make aggressive, essentially unverifiable assumptions about the extent to which the individual members of society collectively value the benefits associated with the quality of life factor in question, say atmospheric decarbonization in this case. When the quality of life factors have been quantified in this way, the economists then model all their inputs against different resource allocations to see which allocation is identified by the model as the one that will maximize social welfare. This resource allocation is then said to be best for all members of society and will be imposed by the policymaker, overriding and adjusting any different resource allocations that may have been made by private ordering arrangements.

Because there is no way to verify the assumptions made about the value that people place on the benefits associated with quality of life modeling inputs like decarbonization, there is no way to show that social welfare modeling outcomes represent anything more than the personal policy preferences of the officials who ran the model. Social welfare modeling can be an interesting fiction that lends undeserved intellectual legitimacy to decisions actually made for the reasons so tellingly revealed by public choice theory and related research.

The American people have many different policy preferences, and they value things like atmospheric decarbonization very differently. In a constitutional republic, the efficient and effective and lawful way for all of those different viewpoints to be filtered and mediated and aggregated is through the actions of the elected representatives of the people in Congress, not through the actions of unelected and unaccountable policymakers issuing top down directives from their positions of unchecked power in executive branch regulatory agencies.

Prime Minister Margaret Thatcher once pointed out that there is no such thing as society. Society is not a collective entity; it is a free association of sovereign individual citizens who met in convention to constitute a government to which they carefully delegated certain enumerated powers.

The concept of society as a thing—an entity whose welfare can and should be maximized by the directives of government policymakers—is just one of the progressives’ abstractions of aggregation that are used to obscure the extent to which top down policy directives disrupt private ordering arrangements to redistribute income by picking winners and losers. Professor Lambert defines regulation to exclude policies that are “solely” concerned with redistribution and claims that his approach to social welfare maximization leaves the subject of regulating to produce equitable outcomes largely “untouched.” His claim is unpersuasive. It ignores the significant extent to which all regulatory directives have a redistributive effect because they disrupt private ordering arrangements. And they all touch on the subject of equitable outcomes because they are driven by the top down presumptions of the “wise few” about the optimal allocation of society’s resources. The only way to optimize the allocation of society’s resources efficiently, and lawfully, is to respect the private ordering arrangements of the individuals who own those resources.

III. A Truly Better Approach to Regulating

Considering the importance of the subject and the author’s truly impressive professional credentials, it is disappointing that the analysis in Professor Lambert’s book is so constrained by his acceptance and use of the progressive concepts and assumptions that have defined the regulatory system for more than 100 years. Professor Lambert tries to improve incrementally the functioning of the current system, judging its performance against standards laid down by progressives for an administrative state. What America really needs is a fundamental and thoroughgoing critique of the current system, judging its performance against standards laid down by the Founders for a constitutional republic governed by the rule of law. That could make the regulatory system truly better.

We like to say that our government is ruled by law and not the passions of men. We like to say that no one is above the law and all are equal before the law. We like to say that we live in a nation governed by the rule of law. In that case, what is required for the regulatory system to be lawful? What are the attributes of lawfulness that we can use to evaluate the current system and to support our efforts to make that system better?

Regulations have the full force and effect of law, the same as the laws enacted by Congress. In fact, regulators enact many more laws than Congress does. In 2016, for example, federal regulators enacted eighteen times the number of laws passed by Congress. Therefore, it is vitally important for regulations, and the procedures that are followed to enact and apply them, to be lawful.

One necessary attribute of a lawful regulatory system is full statutory compliance. To be lawful, regulations as well as the procedures followed to enact and apply them must be in complete

---

50 Lambert, supra note 3, at 5, 256.
51 Clyde Wayne Crews, Jr., Ten Thousand Commandments 2018 72 (Competitive Enterprise Institute, April 2018).
and consistent compliance with all applicable congressional statutes and the controlling court decisions that have interpreted those statutes. To the extent that components and characteristics of the regulatory systems do not meet this statutory standard, then legislative and executive branch leaders should formulate the substantive and procedural improvements needed to bring the system into conformity with the statutory standard.

However, while statutory compliance is necessary for lawfulness, it is not sufficient. To fully satisfy the critically important standard of lawfulness, regulations and the procedures followed to enact and apply them must be in complete and consistent compliance with the fundamental law of our republic embodied in the Constitution, with its provisions interpreted to faithfully conform to their original public meaning. Justice Gorsuch has written eloquently about our constitutional government and the Founders’ rationale for its text and structure.

In enlightenment [political] theory and hard won experience under a tyrannical king the founders found proof of the wisdom of a government of separated powers. In the avowedly political legislature, the framers endowed the people’s representatives with the authority to prescribe new rules of general applicability prospectively. In the executive, they placed the task of ensuring that the legislature’s rules are faithfully executed in the hands of a single person also responsive to the people. And in the judiciary, they charged individuals insulated from political pressures with the job of interpreting the law and applying it retroactively to resolve past disputes. This allocation of different sorts of powers to different sorts of decisionmakers was no accident. To adapt the law to changing circumstances, the founders thought, the collective wisdom of the people’s [elected] representatives is needed. To faithfully execute the laws often demands the sort of vigor hard to find in management by committee. And to resolve cases and controversies over past events calls for neutral decisionmakers who will apply the law as it is, not as they wish it to be.

The founders considered the separation of powers a vital guard against governmental encroachment of the people’s liberties, including all of those later enumerated in the Bill of Rights. A government of diffuse powers is less capable of invading the liberties of the people.52

Through the Constitution, the Founders established a government of limited authority granted by a sovereign people, its enumerated powers checked and balanced by a clear separation of functions, all in order to secure the natural inalienable rights of the people so forcefully affirmed in the second paragraph of the Declaration of Independence. The Founders carefully enumerated the powers and separated the functions of the government in an effort to protect against the tyrannical excesses that always result from the concentration of unbounded power in a single government authority. And through this Constitution, the Founders established the rule of law in our country.

How did our Constitution establish the rule of law? Over the decades, many scholars, jurists, and other commentators have written about the rule of law, its essential attributes, and the significant benefits that can result from the incorporation of those attributes into a country’s system of governance.53 Commentators are agreed that the mere existence of laws and the institutions commonly associated with a legal system do not prove the presence of the rule of law. Countless laws have been promulgated and enforced throughout history, often arbitrarily by abusive tyrants. Dictatorial systems of governance have often included written constitutions, legislatures, and courts; these are the trappings—not the essence—of the rule of law.

The rule of law can be said to exist only if the laws and legal institutions of a country operate to effectively restrain the arbitrary exercise of unbounded authority. Commentators are generally agreed that for a system of governance to embody the rule of law, it must have the following attributes:

- The functions of legislation, executive administration, and adjudication are carried out by clearly separate branches of government, each operating transparently and with essential independence within its respective sphere of responsibility according to the constitution of the country.
- In the legislature, the people’s elected representatives are endowed by the constitution with the authority to enact new rules for general and prospective application. Once enacted by the legislature, the rules remain fixed unless and until amended by the legislature.
- The executive is also answerable to the people and charged by the constitution with the job of faithfully executing and administering the rules that have been enacted by the legislature.
- The judiciary is insulated from politics and political pressures, and it is charged by the constitution with the job of interpreting the rules that have been enacted by the legislature and applying them retroactively to resolve past disputes.
- The laws and the legal system of the country are consistent with the fundamental law contained in its constitution.
- The laws of the country are readily accessible and intelligible to the persons covered by them.
- The legal system treats like cases in like manner, faithfully applying the laws as written, without exception or waiver, to adjudicate questions of right and liability.
- The legal system recognizes and respects fundamental human rights, including property rights, and protects those rights from private parties and public officials who would use arbitrary power to abuse them.54

If these are the essential attributes of the rule of law, two things are abundantly clear: the system established by our Constitution, as eloquently described by Justice Gorsuch, clearly embodies the rule of law, and the current regulatory system does not. Even a cursory evaluation of the current system shows the dramatic extent to which regulation has forsaken the rule of law as described in the list above. This is one of the main reasons many Americans report to pollsters that they think there is “too much” regulation.55

Much of what is concerning about the current system—the consolidation of once separate government functions within a largely unaccountable “fourth branch,” the constantly changing rules, the rules written to pick winners and losers, the endless maze of permitting, the overbearing administration, the politicized enforcement and shakedown settlements, the retroactivity, the waivers, the abiding uncertainty that results from the broadly defined discretionary authority of individual regulators—all of these problems are directly traceable to the lawlessness of the current regulatory system. The essence of the rule of law is the restraint of arbitrary power; the essence of the current regulatory system is the exercise of arbitrary power.

It took years for the current situation to develop. It will take years more, and a great deal of hard work by many dedicated people, to recover a regulatory system based on and faithful to the rule of law and the Constitution. At least the attributes of regulatory lawfulness and statutory and constitutional consistency are clearly defined and well known. Those attributes can be used to identify the specific reforms that need to be implemented to make the regulatory system more lawful and, as a result, better.

A growing body of research increasingly supports the conclusion that a system of governance based upon the rule of law can effectively support the desirable objectives sought by the current regulatory system—things like clean water and a safe workplace—while protecting individual liberty and limiting the threat of government overreach inherent in the progressive system of governance.56 A system of governance or regulation based on the rule of law attains its policy objectives by proscribing actions that are inconsistent with those objectives. For example, this type of regulation would prohibit a regulated party from discharging a pollutant in any amount greater than the limiting amount specified in the regulation. Under this proscriptive approach to regulation, any and all actions not specifically prohibited are permitted.

Hayek and others have noted the significant economic benefits that flow from a system of governance based on rules designed to restrain the arbitrary exercise of power. Such rules make it possible to foresee with fair certainty how power will be applied in given circumstances and, within the open spaces established by such rules, to freely plan one’s own affairs. Such rules encourage individuals to undertake the risks and pursue the rewards associated with productive economic activity and innovation because they know that policymakers will be prevented from disrupting their private ordering arrangements through arbitrary top down directives. And such rules facilitate the efficient allocation of scarce resources because they free individuals to act in ways best suited to the circumstances of time and place best known to each of them.57

While prosperity, progress, and efficiency are significant positive attributes of a system of governance based on the rule of law, they are not the most important. By far, the most important attribute of such a system is a genuine institutionalized respect for the natural rights of each sovereign individual citizen, and a genuine institutionalized commitment to protect the political and economic liberties necessary to exercise those natural rights. All regulations should be based on this kind of respect for the individual. Policymakers should not approach their task as physicians—as superior technocrats working to correct the errors of their inferior fellows in order to improve the health and welfare of a fictitious collective abstraction called society. Rather, policymakers should approach their task as humble public servants charged first with respecting and protecting the natural rights of the individual citizens they work for. Policymakers should review the private arrangements of their fellow citizens with the genuine respect and deference they deserve. And policymakers should never take any action to disturb the private ordering arrangements of their fellow citizens unless that action is clearly authorized by a statute enacted in Congress by the elected representatives of those citizens.

At one point in his book, Professor Lambert quotes a passage from Mere Christianity, by C.S. Lewis, to argue for the originality of How to Regulate. It is appropriate to conclude this review of Professor Lambert’s book with another quotation from Mere Christianity, which forcefully highlights the critical importance of always regulating with genuine respect for the individual and his natural rights, and with a real understanding that the individual citizen is the center of things, not the expert or the thing called society:

If individuals live only seventy years, then a state, or a nation, or a civilization, which may last a thousand years, is more important than an individual. But, if Christianity is true [and human beings are immortal], then the individual is not only more important but incomparably more important, for he is everlasting and the life of the state or civilization, compared with his, is only a moment.58

---


56 See, e.g., Ferguson, Civilization, supra note 53; Deirdre N. McCloskey, Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World (2016).


**Capital Punishment: A One-Sided Contribution to a Complex Debate**

*by John G. Malcolm*

**A Review of:**

*End of Its Rope: How Killing the Death Penalty Can Revive Criminal Justice*,
by Brandon L. Garrett


**Note from the Editor:**

This review discusses a book about the death penalty's decline and criticizes some of its assumptions about the death penalty. 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.


---

**About the Author:**

John G. Malcolm is the Vice President of the Institute for Constitutional Government and the Director of and Ed Gilbertson and Sherry Lindberg Gilbertson Senior Legal Fellow in the Edwin Meese III Center for Legal and Judicial Studies at The Heritage Foundation.

At least you know where he stands right from the start. Brandon Garrett, the inaugural L. Neil Williams, Jr. Professor of Law at Duke University School of Law, begins his book *End of Its Rope: How Killing the Death Penalty Can Revive Criminal Justice* as follows:

We can abolish the death penalty. We must abolish the death penalty. Ten years ago, that declaration would have been laughable, just another liberal fantasy. But no more.

The death penalty in the United States is at the end of its rope. We can abolish it not in a matter of generations, but in a matter of years. And it is imperative that we do so, for its abolition will be a catalyst for reforming our criminal justice system.

What follows is, not so much a liberal fantasy, but a lengthy, one-sided elaboration of the arguments that liberals and other death penalty opponents (some of whom are conservatives) have been making for years. This is a shame.

While Garrett certainly does an admirable job of laying out his side's perspective, the death penalty is a contentious and complicated issue with strong arguments to be made on both sides. The topic is deserving of a more thorough, nuanced, and balanced treatment than it receives here. Indeed, in arguing that the death penalty is unjust and unconstitutional, Garrett makes repeated references to the Eighth Amendment's prohibition against "cruel and unusual punishment." He does not, however, mention the fact that, whether it is good or bad policy, there are several explicit references in the Constitution itself condoning the use of the death penalty; specifically, the Fifth and Fourteenth Amendments provide that "No person shall be held to answer for a capital . . . crime, unless" indicted by a grand jury, that a person cannot twice "be put in jeopardy of life . . . " for the same offense, and that a person may not be "deprived of life . . . without due process of law."

Garrett claims that the death penalty has declined in usage and popularity in recent years. While he is certainly correct that there have been fewer executions recently, it is far less clear that public support for capital punishment has significantly waned.

The death penalty is still favored by a majority of Americans. According to a June 2018 poll by the Pew Research Center, 54% of Americans favor capital punishment for those convicted of murder, up 5% over the last two years. 2 These numbers are similar to those found in another poll released by Quinnipiac in March, which showed that Americans support the death penalty for people convicted of murder by a margin of 58% to 33%, with 9% undecided, and that 64% of Americans feel even more strongly that the death penalty should not be abolished nationwide. 3 Death penalty opponents, including Garrett, are quick to note that support for the death penalty is down quite

---

1 He was the Justice Thurgood Marshall Distinguished Professor of Law at the University of Virginia School of Law when this book was published.


3 Most U.S. Voters Back Life Over Death Penalty, Quinnipiac University National Poll Finds; Voters Back Anti-Gun March 2-1, But Say It Won't
a bit from its high-water mark of 80% in 1994, but they fail to mention it is way up from the 42% support it enjoyed in 1966.\(^4\)

Moreover, when presented with the facts in individual cases—such as those of Timothy McVeigh, the Oklahoma City bomber who killed 168 people and injured over 800 more, or Khalid Sheik Mohammed, the self-professed mastermind of the 9/11 plot that killed nearly 3,000 people—support for the death penalty goes much higher. Indeed, according to an October 2017 Gallup poll, 39% of Americans do not believe that the death penalty is imposed often enough, compared with only 26% who believe it is imposed too often.\(^5\) In 2016, despite a well-funded campaign by death penalty opponents, voters in California, Oklahoma, and Nebraska voted to retain the death penalty.\(^6\) The citizens in those and most other states (31 in total) continue to believe that the death penalty—despite its flaws—is the only punishment befitting those who commit certain particularly heinous and depraved murders.\(^7\)

As Garrett notes, the number of executions has, as a general matter, declined recently, although the number of executions over the last four years (28 in 2015, 20 in 2016, 23 in 2017, and 14 so far in 2018, with several more scheduled before the year ends) exceeds the number carried out from, for example, 1988 to 1991 (11, 16, 23, and 14, respectively).\(^8\) Garrett attributes this decline to several factors, including the facts that the murder rate has dropped precipitously over the last twenty years and that defense attorneys in capital cases are now better funded and better trained—both of which are laudable developments. Nobody supports incompetent defense attorneys, racist prosecutors, or bad judges, especially in capital cases, and Garrett certainly does a thorough job of chronicling seemingly every instance where such bad actors have been involved and the convictions have been overturned (although, in his view, others have not been overturned when they should have been).

While Garrett notes the increased costs associated with capital cases and the ever-increasing length of time between conviction and execution in capital cases, he neglects to discuss the concerted strategy by death penalty opponents to drag out the process and to make it as costly as possible. In other words, death penalty opponents are urged to wage what Justice Samuel Alito has called “a guerilla war against the death penalty,”\(^9\) even in cases in which their perfectly competent clients (such as convicted double murderer Scott Dozier) wish to end their appeals and face execution.\(^10\) Houston attorney Katherine Scardino, who has been referred to as “the Clarence Darrow of death penalty lawyers in Texas,” has a word of advice for anyone appointed to a capital case: “Spend money. That will get everybody’s attention.”\(^11\) And it certainly does. In short, if death penalty opponents cannot persuade their fellow citizens to abolish the death penalty on the merits, they will simply try to bleed the system dry—and they often succeed, as Garrett candidly acknowledges when he points out that “mounting costs may explain why rural counties have almost entirely stopped death sentencing, and why, over the past two decades, death sentencing has retreated to a handful of large, densely populated counties that can still afford it.”

In 1985, the average time between a death sentence and execution was just under 6 years.\(^12\) By 2013, it was 15 years, 6 months (which was actually slightly lower than the previous two years).\(^13\) Two people executed earlier this year (Carlton Michael Gary and Robert Van Hook) had been sentenced to death more than 30 years ago. While there are nearly 750 inmates on death row in California, California has executed only 13 people since 1978, and none in the last decade.\(^14\) No wonder many prosecutors and citizens decide that, even though they want to retain the death penalty, the game is not worth the cost or the candle.

Garrett also points to racial disparities in death penalty cases and boldly asserts that there is only one explanation for these disparities: endemic racism. That explanation may or may not be true, but Garrett does not even attempt to probe alternative possibilities. It is certainly true that, according to the last census,

---

African-Americans constitute roughly 12.3% of the population in this country, but, as of 2016, 42.3% of the population on death row. This is a shocking disparity until one considers that, as of 2016, African-Americans constitute 52.6% of those arrested on charges of murder and non-negligent manslaughter and only 34.3% of those who have been executed since the death penalty was reinstated in 1976. Moreover, in terms of inter-racial murders, there are far more black-on-white crimes than there are white-on-black crimes. Inter-racial violent crimes are committed at much higher rates by strangers usually in the act of committing other crimes such as a rape or armed robbery, the kinds of aggravating factors that are often taken into account by prosecutors in deciding whether to pursue the death penalty. None of this is to deny the possibility that race plays a role in decisions about when to seek or impose the death penalty. But these facts are offered to suggest that there may be other, non-racial explanations for some of these disparities; Garrett does not mention, much less discuss, any such alternative possibilities.

A reader would think based on the descriptions in this book that every time the death penalty is imposed, it is due to ignorance or heartlessness by juries, crooked and overzealous prosecutors, biased “Hang ‘Em High” judges, incompetent defense attorneys, and poor and misunderstood defendants. In leaving that impression, Garrett mischaracterizes the process, the difficult decisions that jurors face, and the sobriety and earnestness with which they approach their grim task.

There are other noteworthy omissions. For example, Garrett points to a handful of horrific, high-profile botched executions as another reason why some people have turned against the death penalty. He fails to mention, however, that the successful lobbying efforts of death penalty opponents—which have resulted in reputable drug manufacturers refusing to supply drugs for lethal injections and in trained medical professionals refusing to participate in carrying out executions—have dramatically increased the likelihood that executions will be botched.

And Garrett points to a number of so-called exonerations (some of which are actual exonerations in the sense that it was definitively determined that the accused did not commit the murder, while others are reversals because of some procedural irregularity or new evidence that casts doubt upon the verdict) as proof that innocent people have clearly been executed since the Supreme Court reinstated the death penalty in 1976. Despite the best efforts of death penalty opponents, however, it has never been definitively established that anyone has been wrongly executed since then. But this does not stop Garrett from confidently declaring that “[d]eath penalty states are no doubt still executing innocent people.” Has an innocent person been executed since the Supreme Court reinstated the death penalty? It is certainly possible. As is the case with any human endeavor, mistakes can be made in the imposition of the death penalty. Nonetheless, in addition to the fact that there would still be arguments to support the death penalty even if an innocent person has been executed, it is also true that virtually all death penalty cases receive multiple layers of judicial review. This is especially so in cases when there is even a colorable claim of innocence, and governors have not hesitated to commute death sentences, even when the accused’s guilt was never in doubt, when they believe that imposing the sentence would constitute a miscarriage of justice.

But Garrett’s agenda is more radical than simply abolishing the death penalty. He makes it quite clear that life without parole is equally objectionable, if not more so, in his eyes because “life rows have mushroomed in size, dwarfing the population of death rows even at their height.” Instead, he favors a justice system based on mercy which is, in turn, premised on “empathy for another person,” specifically, the perpetrator of crimes. And not just non-violent crimes. According to Garrett, we “have to embrace mercy for the most serious offenses,” and ought to be “willing to shorten prison terms and release” those who commit those offenses. After all, Garrett declares, “This is the land of the free.” Under this utopian (some might say Pollyannaish) view of the world, those who have committed heinous crimes will simply be overwhelmed by this gesture of grace, see the light, and go forth into the world and sin no more. But this is a mighty bold and risky gamble in a world where recidivism rates among formerly-incarcerated individuals remain staggeringly high.

In making his case, Garrett implies that in order to be in favor of criminal justice reform, one must be against the death penalty (“the sudden decline in the American death penalty is a social trend that speaks volumes about the present and future of our criminal justice system”) and that one will simply not be able to address the former unless and until the latter is abolished (“the death penalty’s demise will allow us to focus on remedying the myriad of problems with the current criminal justice system). As a supporter of much of the criminal justice reform movement,
color me skeptical. One can focus on remedying the lingering problems with the death penalty without abolishing it, while at the same time addressing some of the problems with our criminal justice system. Indeed, working to remedy some of the problems with death penalty procedures will likely help address some of the problems with the broader criminal justice system.

None of this is to suggest that Garrett does not point to some very legitimate problems with the current death penalty process and the broader criminal justice system. He points to police interrogation and suspect identification techniques that may be unduly coercive or suggestive, and which may result in false confessions and improper out-of-court identifications. Others have also pointed out problems with these techniques.22 And Garrett points to problems with forensics labs, which can be particularly troublesome given the outsized influence that forensic evidence, with its patina of objectivity and irrefutability, can have in the courtroom (a problem that I have also written about).23 These are certainly areas that call for improvement, but they hardly support the call for abolition of the death penalty on their own.

He also points to some promising developments on the criminal justice reform horizon. For example, Garrett notes the increasing availability of mental health courts which are designed to deal with certain offenders who suffer from severe, untreated mental illnesses that likely precipitated the crimes they committed. These and other diversionary courts, such as drug courts and veterans courts, may reduce recidivism and constitute a more just way of treating certain categories of offenders (whether that should include murderers, of course, is a different matter).

Is the death penalty on its way out? Who knows? Are there sound reasons to support or oppose the death penalty? Of course. This is a serious subject about which reasonable people can and do disagree, and disagree passionately. If you want to join that debate and are looking for an effective opening argument against the death penalty, then this book is for you. If, on the other hand, you are looking for a balanced exposition of a complicated and contentious issue, then keep looking.


Can Americans Reconcile Our Constitutional System With an Expansive Administrative State?

by Ted Hirt


Note from the Editor:

This review discusses the history of the tension between constitutional and administrative government in the United States.

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.


In Bureaucracy in America: The Administrative State’s Challenge to Constitutional Government, Joseph Postell, Assistant Professor of Political Science at the University of Colorado-Colorado Springs, provides a detailed and scholarly review of our government’s reliance on administrative agencies to administer benefits and regulate citizen conduct. The story he recounts is fascinating, and it manages to be both detailed and engaging.

Postell develops several themes. First, he argues persuasively “that there always has been a tension between administrative power and American constitutionalism.”1 Tracing the history of the administrative state from the founding to today, Postell challenges the thesis of some scholars that the administrative state’s “crisis of legitimacy” is simply a recent phenomenon that has been ginned up by conservative critics.2 Second, Postell contends that attempts by various progressives and modern liberals to incorporate principles of American constitutionalism into the administrative process has not resolved this tension in a satisfactory manner.3 Postell next examines the embrace of deferential judicial review of agency action by some conservative jurists. Finally, he asks how constitutional government can be rescued from an expansive administrative state.4

Postell’s work focuses predominantly on our nation’s constitutional history and how administrative agencies have functioned within it. That context helps us understand the current crisis of legitimacy of the administrative state. Recounting the past also provides some clues to how our republic can reconcile the competing values of constitutionally required separation of powers, limited government, and government accountability on one hand, and efficiency in administration on the other.

I. FROM THE COLONIAL PERIOD TO THE CONSTITUTION

One might expect a history of the American administrative state to begin in 1787 with the Constitution’s deliberately formulated structure of legislative, executive, and judicial powers, articulated in Articles I, II, and III of that basic charter. But Postell emphasizes that the Framers, by the time they reached the 1787 Philadelphia Convention, already would have been familiar with the exercise of administrative powers during the colonial period and at the state and national levels during and after the American Revolution.5 These “lessons of experience” informed the debates and decisions that resulted in the Constitution.6

While readers are likely familiar with the rivalry between colonial legislatures and royal governors—in which legislatures asserted their own rights to tax and spend, contrary to assertions of top-down control by the British Crown—few may know that local courts were an important locus of colonial government...
Postell explains that justices of the peace functioned not only as agents of their local communities, but as one-man administrative agencies. Justices of the peace regulated a wide range of conduct, including building highways, establishing inns and liquor retailers, and some food sales. Postell emphasizes that, despite our current controversies, administrative law “did not start with judicial review of administrative activity; the courts were the administrators themselves.” And all local officials— including justices of the peace—were accountable to the citizens through frequent elections.

Judicial activity in the colonial era also established a “basic principle” of American constitutionalism: that executive officers “could not be placed above liability for following the law.” For example, common law damages actions were available against sheriffs and jailers for wrongful conduct, and the courts refused to enforce writs of assistance, which were general warrants sought by British customs officers to search warehouses and other private property for illegal or untaxed goods. Two other factors weakened administration and made it more accountable: the “absence of a powerful executive and administrative branch weakened the ability of the government to enforce law against the wishes of the community,” and that phenomenon ensured that law was consistent with custom, a principle “central” to the common law.

The advent of American independence led the states to experiment with new ways of exercising power. Postell notes that the revolutionary government’s early experiment with Committees of Safety fell into disfavor because the Committees tried to exercise legislative and administrative power independently of state legislatures. In contrast, the Continental Congress created ad hoc committees to address an array of war-related functions. The resulting “administrative sprawl” was not only inefficient, but it also mixed legislative and administrative functions. When the Continental Congress established multi-member boards and committees to help conduct the war effort (raising revenue, securing munitions and supplies, etc.), it encountered a different problem: members were unaccountable. But single-member boards did not pose that problem because the “concentration of accountability in a single person” and the continuity provided by that person’s leadership ensured “greater energy, efficiency, and responsibility to the ends set forth by the authorizing legislature.”

Postell observes that several principles of administrative constitutionalism emerged from these experiences: the fact that judges could be relied upon as agents of regulation and administration, the need for administrators to be accountable to the people, the usefulness of judicial review and legal checks to guard against arbitrary administrative actions, the importance of separating administrative power from legislative interference, and the wisdom of “unitary executive structures.” To support this thesis, Postell cites portions of debates from the Philadelphia Convention, state ratification conventions, and the Federalist Papers. Those debates make it clear that the Founders opted for a strong principle of non-delegation of legislative authority to the executive. The Founders also rejected the notion of a Council of State that would constitute a non-unitary executive (analogous to the multi-member boards the Continental Congress rejected). James Madison and Alexander Hamilton defined republicanism as a system in which government had to have an “immediate” relationship to the people, rendering government officials accountable to the latter. In Federalist No. 52, Madison stated that it was “essential to liberty” that the government should have an “immediate dependence on, & an intimate sympathy with the people.” Postell contends that legislation by administrative agencies is inconsistent with that principle insofar as administrators are not directly (or even indirectly) elected by the people. Finally, the separation of powers was a key principle undergirding the Founders’ vision for government; one consequence of this is that administrative agencies should not be able to exercise multiple or “blended” powers within their organizations.

II. Administrative Power in the Early Republic

After the Revolution, state and local governments continued to develop various means of deploying administrative power. For example, governments developed systems of inspection and licensing of commodities and containers, and they regulated entry into various occupations and professions. The governments also regulated common carriers, chartered entities such as banks, and transportation companies such as ferries. The Founding generation accepted as legitimate a range of regulation; they were
not laissez-faire ideologues. But they cared about how those exercises of regulatory authority were administered and, in turn, made accountable.

In most states, authority over local issues was transferred by the state government to the relevant local governments. As a result, county administrative officers held much of the administrative power that would otherwise have resided with state officials. These local officials were more likely to be directly elected or appointed by the legislature than they were to be appointed by the executive. Local officials could enforce some laws through penalties imposed by an administrative officer, but more substantial penalties were dispensed in court systems. Local boards of health and sanitation were created, but with circumscribed authority. City councils created boards to investigate discrete social problems, such as health hazards and sanitary practices, and report back to them. There was some blending of lawmaking, enforcement, and judicial powers, but the activities were local and did not pose a risk of widespread abuses of power. State legislatures were disinclined to delegate authority to administrators, so they often enacted elaborately detailed specifications for how administrators were to carry out their tasks. Courts continued to act as forums for private citizen enforcement of the laws and for review of the legality of administrative agency action. The virtues of administration during this period included a "constrained administrative apparatus that was tightly bound to public opinion and promoted self-government at the local level," and that the system avoided the "creation of an elite bureaucracy" that would have been removed from public opinion or oversight or unaccountable to the court system.

At the national level, Postell says, Congress “adhered to a nondelegation principle” and was “vigilant in retaining the power to make the law.” Congress legislated the routes of post roads, specified the locations of lighthouses, and the Interstate Commerce Act, which created an administrative inspection system, was a precursor to the modern independent regulatory commission. Postell argues that, in that statute, Congress created very limited agency authority that cannot be analogized to the “expansive rulemaking powers of the modern administrative state.”

Also during this antebellum period, the federal government maintained the principle of the unitary executive that placed all administrative authority under presidential control. Washington and his successor presidents generally adhered to this theory of administrative power. And the president’s power to remove subordinate officials—a power not encumbered by Congress—also was generally accepted.

Courts served as enforcement mechanisms for legislation, rather than as administrative agencies as in the colonial era. From reviewing Supreme Court precedent, Postell finds continuity in judicial review of executive action from the Founding to the antebellum period. The Supreme Court maintained the principle of de novo review of agency action, although Chief Justice Roger B. Taney suggested that some deference to agency decision-making might be appropriate. But, because cases challenging agency actions usually involved requests for the extraordinary writ of mandamus, which courts are disinclined to grant when the issue involves discretionary agency acts, it is not clear if broader principles of agency deference can be derived from those court decisions.

III. Post-Civil War and Progressive Era Developments

Postell devotes much of his book to describing various efforts to create administrative agencies after the Civil War. Some commentators “mark the 1880s” as the decade in which the administrative state was born. The 1883 enactment of the Pendleton Act, which created a competitive civil service, and the 1887 enactment of the Interstate Commerce Act, which created the Interstate Commerce Commission (ICC), are the prominent regulatory landmarks of the post-Civil War period cited for that

29 Id. at 62.
30 Id. at 64.
31 Id. at 63.
32 Id. at 65.
33 Id. at 66.
34 Id. at 65.
35 Id. at 66.
36 Id. at 67-69.
37 Id. at 69-71.
38 Id. at 72.
39 Id. at 75.
40 Id. at 76, 78.
41 Id. at 96-102.
42 Id. at 101.
43 Id. at 81.
44 Id. at 81-82.
45 Id. at 84-89.
46 Id. at 89-91.
47 Id. at 117-24.
48 Id. at 117.
49 Id. at 120.
50 Id. at 120-21.
51 It is not clear why Postell does not address legal developments during the Civil War other than to note, in passing, the creation of the Department of Agriculture in 1862. Id. at 129. It would make sense, for example, for Postell to examine the National Bank Act of 1862 or other expansions of national power that occurred in conjunction with President Lincoln’s and Congress’ pursuit of victory in the war against the Confederacy.
52 Id. at 127.
Commission could appoint individuals who were not accountable
enforcement suit, and the ICC’s determinations would be
the agency could ask the United States Attorney to file a civil
railroads, but the ICC itself did not have enforcement powers;
to bring common law suits. It authorized the ICC to investigate
should simply defer to its factual determinations. 65 Courts also
in judicial proceedings, rejecting the agency’s argument that courts
addition, the courts permitted carriers to introduce new evidence
enforce these declarations.62 Postell contends that the authority
unjust and unreasonable—it had to rely on the courts to actually
competitive practices, but the ICC only could
granted to the ICC was not an “open-ended grant of discretion,”
insofar as the phrase “reasonable and just” had a well-defined
meaning.63 Over time, the ICC acquired authority to issue cease
and desist orders, but that authority was not self-executing.64 In
addition, the courts permitted carriers to introduce new evidence
in judicial proceedings, rejecting the agency’s argument that courts
should simply defer to its factual determinations.65 Courts also
determined that the ICC could only issue judgments about past
rates; it could not engage in prospective ratemaking.66 From this
history, Postell concludes that the ICC “was not intended to be,
nor was it in its initial practice, a powerful independent regulatory
commission in the progressive sense.”67 The period immediately
following the Civil War, then, contrary to some popular belief,
did not birth the administrative state as we know it today.

The Progressive Era, however, ushered in a sea change in
administrative law and American constitutionalism, founded in
part on the advocacy of reformers and like-minded scholars for
fundamental reforms to the American political system, including
elimination of the indirect election of Senators and institution
of the initiative, referendum, and recall as democratizing
measures.68 The Progressives questioned the continued wisdom of
a government conducted by elected representatives; Professor
Herbert Croly, a prominent Progressive, believed that the United
States should become a “more highly socialized democracy” under
which there would be an “efficient national organization.”69 Other
Progressive thinkers concluded that the Constitution’s tripartite
separation of powers was unworkable or outmoded.70 These
thinkers advocated a consolidation of powers in administrative
departments, in which experts could devise and implement policy.71

President Theodore Roosevelt embraced the Progressive
movement,72 and Congress enacted, among other things, the
Hepburn Act of 1906, granting ratemaking and final adjudicatory
powers to the ICC.73 The Supreme Court explained that it
would defer to the agency’s factual determinations, but that
it would decide legal issues de novo.74 President Woodrow
Wilson and some of his advisors, however, did not embrace the
centralization of government envisioned by President Roosevelt’s
New Nationalism, under which industrial monopolies would be
permitted but intensely regulated.75 Louis Brandeis, a prominent
lawyer and scholar before he became a Supreme Court Justice,
was skeptical of the use of executive government power to regulate
the economy or big business, preferring the use of court-based
enforcement of statutes like the Sherman Anti-Trust Act.76 In
the Federal Trade Commission Act of 1914, Congress authorized
the new Federal Trade Commission (FTC) to define what constituted
an “unfair method of competition or deceptive act or practice,”

53 Id. at 127, 136-62.
54 Id. at 127.
55 Id. at 127-29.
56 Id. at 137, 144.
57 Id. at 137-41.
58 Id. at 142-43.
59 Id. at 143.
60 Id. at 150-52.
61 Id.
62 Id. at 153.
63 Id. at 154-55.
64 Id. at 154.
65 Id. at 162-63.
66 Id. at 163.
67 Id. at 164.
68 Id. at 167, 169-71.
69 Id. at 172-73.
70 Id. at 174-76.
71 Id. at 174-78.
72 Id. at 194-96.
73 Id. at 185-187.
74 Id. at 188 (citing ICC v. Illinois Central Railroad Co., 215 U.S. 452, 470 (1910) and ICC v. Northern Pacific Railroad Co., 216 U.S. 538, 543-44 (1910)).
75 Id. at 196-200.
76 Id. at 199-200.
but the FTC had to go to court to enforce that provision.77 Postell discerns in this period some increase in judicial deference to agency decisions.78

IV. The New Deal and the Administrative Procedure Act

Although the administrative state established a foothold in the early twentieth century through the establishment of a few agencies like the FTC and the Federal Power Commission, President Franklin D. Roosevelt’s New Deal tried to expand the administrative state “dramatically and to insulate its decisions even further from judicial oversight.”79 The New Deal “introduced a plethora of regulatory programs, each delegating broad powers to a newly-created or existing administrative body.”80 But the path to an expansive administrative state held significant obstacles, and the battles over the limits of executive agency power still resonate today.

For example, in 1935, the Supreme 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.81 In the well-known case of A.L.A. Schechter Poultry Corporation v. United States,82 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” constituted an improper delegation of legislative powers.83 In another case, the Court determined that the president’s otherwise broad removal powers did not apply to members of the FTC, explaining that FTC members, unlike cabinet officers, have quasi-judicial and quasi-legislative duties, and that such expert duties had to be exercised independent of the president.84 Postell observes that the Court’s reasoning here embraced the Progressive vision of independent experts operating in an environment uncoupled from traditional notions of separation of powers.85 Some people in the Roosevelt Administration invoked constitutional language in rejecting the proposed establishment of agencies that would be independent of the President.86 But the Reorganization Act of 1939, which gave President Roosevelt the authority to reorganize the executive branch, exempted the most important regulatory agencies from his reorganization authority.87

Eventually, the Supreme Court did not stand in the way of important New Deal initiatives or other social reform legislation.88 For example, in National Labor Relations Board v. Hearst Publications,89 the Court deferred to the National Labor Relations Board’s (NLRB) determination as to whether specific workers qualified as “employees” for purposes of the National Labor Relations Act.90 Justice Rutledge opined that, although the Court would resolve questions of statutory interpretation, the Court would not substitute “its own inferences of fact” for the NLRB’s.91 The Court explained that “where the question is one of specific application of a broad statutory term in a proceeding in which the agency administering the statute must determine it initially, the reviewing court’s function is limited.”92

Two other Supreme Court-related developments in this era deserve mention as part of the context for the emerging administrative state. First, there was some expansion in the deference that courts gave to the agencies. For example, in Crowell v. Benson,93 Chief Justice Charles Evans Hughes opined that Congress could give administrative agencies the power to make final determinations of fact in cases in which individual rights were adjudicated.94 Second, some New Deal theorists thought they could use of the principle of standing, under which judicial review is available only to litigants who can demonstrate a specific injury to their interests resulting from the challenged government action,95 to reduce judicial review of administrative actions.96

In Ashwander v. Tennessee Valley Authority,97 Justice Brandeis asserted in his concurrence that the Court had exercised caution in reviewing the validity of Acts of Congress, and had “restricted exercise of this function by rigid insistence that the jurisdiction of federal courts is limited to actual cases and controversies; and that they have no power to give advisory opinions.”98

Congress enacted the Administrative Procedure Act (APA) in 1946 because of dissatisfaction with the New Deal’s vision of an expansive administrative state unchecked by judges. Readers may

---

77 Id. at 202-03.
78 Id. at 204-05.
79 Id. at 207.
80 Id. at 208.
82 295 U.S. 495 (1935).
83 Id. at 536-42.
85 Postell at 212.
86 Id. at 217-18.
87 Id. at 218.
89 322 U.S. 111 (1944).
90 Id. at 131-32.
91 Id. at 130-31.
92 Id. at 131.
93 285 U.S. 22 (1932).
94 Id. at 53-55.
96 Postell at 220-21.
98 Id. at 345-46.
be surprised to learn that prominent lawyers expressed objections to the new administrative state even at that time. In 1933, the American Bar Association’s Special Committee on Administrative Law advocated transferring judicial power from administrative agencies to independent tribunals such as the courts.99 Roscoe Pound, a Dean of Harvard Law School who had been a progressive legal theorist, spearheaded the ABA’s legislative reform efforts, which culminated in the Walter-Logan Act in 1939, a bill that would have established independent review boards and provided substantial judicial review in suits challenging agency action.100 President Roosevelt vetoed that measure, but its proponents challenged several Progressive and New Deal assumptions about administrative agencies.101 These proponents of legislative reform wanted a separation of administrative powers from judicial powers within agencies, procedural checks on the powers of agencies that could be enforced by the courts, and more robust review by the courts of agency decisions.102

The 1946 debates on the APA reflected continued disagreement about administrative agencies’ exercise of broad powers. Some members of Congress invoked separation of powers principles in denouncing the agencies’ “usurpation” of legislative powers, and others asserted that the agencies improperly wielded judicial authority.103 Other critics of agency power focused on the apparent lack of transparency or consistency in agency decisions, which often relied on trial examiners who made initial decisions that were later reviewed by agency heads who had not participated in the underlying trial.104

Ultimately, however, the APA as enacted did not meet the broad objectives of critics of the administrative state. These critics hoped that the APA was just the first step towards reform of the administrative state.105 Postell argues that the debate leading to the passage of the APA demonstrated “a genuine consensus” among members of Congress that the administrative state needed curbs on its powers and that the administrative state “threatened basic principles of constitutional government.”106 But greater reforms did not occur and, Postell contends, there was no pronounced difference in the rigor of judicial review under the new statute.107

V. Liberal and Conservative Reactions to the Modern Administrative State

Liberals, having created the modern administrative state, eventually developed means—including some that relied on the protections of the Constitution—to limit that state’s reach, at least where individual rights were implicated. These reformers tried to use administrative law to control the bureaucracy.108

The newer agencies that were established in the 1960s and 1970s, such as the Environmental Protection Agency and the Consumer Products Safety Commission, differed in focus (albeit not in impact or importance) from the predecessor agencies of the New Deal period. The older agencies like the Securities and Exchange Commission and the NLRB had focused on the regulation of the nation’s economic order, while the newer agencies addressed environmental, consumer, and broader societal issues.109 At the same time, the Supreme Court, responding to arguments that individuals who participated in government programs had procedural rights in those benefits, interpreted the Due Process Clause of the Fourteenth Amendment to require evidentiary hearings before benefits could be terminated.110 The expansion of the administrative state also was accompanied by a reorientation in its goals, away from the Progressive notion that regulatory questions had an “objectively” correct answer, to a more ideologically-laden inquiry that focused on the “fundamental values” that transcended the administrative process.111 At the same time, some reformers saw that agencies were vulnerable to capture by regulated industries, and they contended that part of the solution to that problem was increased judicial review of agency decisions.112

Some courts, especially the United States Court of Appeals for the District of Columbia Circuit, increased their oversight of agencies through expansive interpretations of the rulemaking provisions of the APA, mandating various procedures and scrutinizing agencies’ decisions more rigorously.113 Standing doctrine in the federal courts also was relaxed, easing access to the courts, sometimes by Congress, but often by the courts themselves.114 From these developments, Postell concludes, a new vision of the administrative state emerged, one that emphasized “participatory democracy, judicial oversight, and the increased influence of interest groups in administrative decision making.”115 Under this theory, judges were viewed as “guardians” of the administrative process, and the administrative state also was being democratized, but outside the “traditional representative institutions of the Founders’ Constitution.”116

This vision, however, was unexpectedly disrupted by a “conservative counterrevolution” in which some conservative judges questioned some of the premises of the reformers’ new vision. In doing so, the judges ironically strengthened the power of

99 Postell at 228.
100 Id.
101 Id. at 232-36.
102 Id. at 233-35.
103 Id. at 237.
104 Id. at 238.
105 Id. at 239-40.
106 Id. at 243.
107 Id. at 244-45.
Postell notes that Justice William Rehnquist’s 1978 opinion in *Vermont Yankee Nuclear Power Corporation v. Natural Resources Defense Council,*118 in which the Court rejected the notion that courts should supplement the procedures prescribed in the APA for the informal rulemaking process.119 Postell asserts, however, that agencies nevertheless became more cautious about engaging in rulemakings that could be subjected to judicial review and increasingly turned to less formal means of decision-making such as “interpretative rules” and “statements of policy.”120 Postell does not endorse these methods, which he considers to be an end-run around APA requirements.121

Postell also points out that some conservative judges embraced the *Chevron* doctrine, under which courts defer to an agency’s interpretation of ambiguous congressional language.122 Although the *Chevron* decision was written by Justice John Paul Stevens, Postell observes that Justice Antonin Scalia endorsed *Chevron* to the extent that it represented, in his view, a proper division of authority between the executive and judicial branches.123 Justice Scalia explained his rationale for that position in *United States v. Mead Corporation:* if Congress left an ambiguity in a statute that would be administered by the agency, courts should presume that Congress intended to give the agency discretion, “within the limits of reasonable interpretation,” on how to resolve the ambiguity.124 Congress committed both the enforcement of the statute and the “initial and primary interpretation” of the statute to the agency, not the courts.125 Postell notes, however, that Justice Scalia explained in *Rapanos v. United States* that Congress must legislate with clarity with respect to the boundaries of federal and state authority.126 Postell concludes that conservatives who have defended the *Chevron* doctrine have inadvertently permitted agencies to “update statutes by creative interpretation,” thereby improperly appropriating legislative authority to themselves, and that agencies’ ability to reinterpret their own authority typically has resulted in an expansion of that authority.127 Agencies nimby fill the gaps left by Congress, including by making policies that Congress never enacted.128

Postell points out that Chief Justice Rehnquist and Justice Scalia preferred to empower the executive branch over the judicial branch because they saw agencies—which are part of the executive branch, one of the political branches that are elected by the people—as more accountable to the people than the judiciary.129 Postell rejoins that each branch has a critical balancing role in our constitutional system of separation of powers, arguing that the Founders would not have favored the accumulation of legislative, executive, and judicial powers in one of the political branches, as has happened in some agencies.130 Postell also contends that agencies have been considered by some to be apolitical and thus insulated from the elected president—the source of their supposed accountability to the people.131 The conservative counter-revolution opposing the administrative state has resulted in some cabining of agency authority through the application of administrative law, but the premises of how the administrative state governs have been left undisturbed, thus permitting, for example, broad delegations of legislative power by Congress to agencies.132 And Postell points out that Justice Scalia’s application of a majoritarian approach to the Constitution—by a strict understanding of the standing doctrine—has resulted in some restraint on judicial review.133

VI. What is the Way Forward?

Postell concludes that an inevitable tension exists between American constitutionalism and the administrative state, and he contends that his historical survey demonstrates that the tension has been recognized and debated across the political spectrum since the Founding.134 But he also says that administrative law doctrines do not solve the problem of a proper allocation of powers between administrative agencies and Congress as prescribed under the Constitution.135 Postell places some hope in recent comments by Chief Justice John Roberts and Justice Clarence Thomas that express skepticism about the asserted reach of agency powers.136 Postell rejects the claims of some commentators that concerns about the compatibility of modern administrative government and the core principles of constitutional law have been “gradually overcome or modified out of existence.”137 To Postell, these controversies are very much alive, and should remain so.
This review cannot do justice to the complex history that Postell recounts. That history is both needed and refreshing because Postell offers a competing vision of the place of the modern administrative state in our constitutional system. But, in my judgment, Postell does not sufficiently point lawmakers or jurists in the right direction—he does not do enough to define the appropriate balance between administrative power and checks by Congress and the courts. His goal is simply to explain and develop an historical narrative, not to prescribe solutions to the very difficult problems he describes, but his excellent historical investigation could have yielded more helpful insights for policymakers had he offered them. Postell’s analysis should provoke more debate among scholars about these issues, and a more searching debate by policymakers about the role of administrative agencies in our constitutional system.138

GIVING CREDIT FOR SHAPING THE CONSTITUTION

by Karen Lugo

A Review of:
The Lives of the Constitution: Ten Exceptional Minds That Shaped America’s Supreme Law, by Joseph Tartakovsky

Note from the Editor:
This review discusses a new book about ten people who shaped the Constitution and is both critical and complimentary.
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.


No two historians—professional or amateur—would compile the same list of the top ten influencers of the Constitution, and no doubt every reader will quibble with at least one or two of Tartakovsky’s selections. But whether you agree or disagree with any given selection in the ten, Tartakovsky has inarguably profiled some exceptional minds and some incredible characters. The descriptions of some—like Ida B. Wells, who campaigned indefatigably for African-American criminal defense rights and women’s suffrage, and Stephen Field, who was a rough-and-ready deliverer of justice to gold miners in San Francisco before appointment to the Supreme Court—bring vital elements of personal biography to the historical account.

But Tartakovsky does not always make a satisfying case for how the individuals he profiles shaped the Constitution. He does not tell his readers how he chose the ten, and he does not provide a calculus for measuring constitutional influence. And especially in the cases of James Wilson, Alexander Hamilton, and Robert Jackson, there is more to their histories that would help to explain how they influenced the Constitution. In some cases, the constitutional shaping was arguably destructive, as with Woodrow Wilson. Yet Tartakovsky dodges the questions that undoubtedly will rise in some readers’ minds.

An important emphasis in the book is that, although the Constitution has been pronounced irrelevant at various intervals in American history, these are also the eras that offer crucial lessons for how to ensure its future survival. Tartakovsky concludes that the Constitution must be cherished and that, so long as it is, it “will be displaced no sooner than an ant tips over the Statue of Liberty.” But he does not contemplate the eventuality of failure to cherish it. What can we expect when multiple generations of students have been taught that the Constitution and its framers—that the American project writ large—are fatally flawed? The author points to hopeful constitutional revivals, but he does not account for a time of unprecedented and sustained attack on constitutional government.

This review will proceed by commenting on Tartakovsky’s treatment of eight of the ten lives; I leave out his discussions of two foreign views of the Constitution: those of Alexis de Tocqueville and James Bryce. Some of my comments summarize or elaborate on Tartakovsky’s work. Others consider his profiles and interpretations more critically.

I. JAMES WILSON

The life and work of James Wilson is given richly deserved attention in The Lives. This is coincident with the efforts of Professor Hadley Arkes, who founded the James Wilson Institute on Natural Rights and the American Founding to help law students, scholars, and practitioners discover the contributions of James Wilson and other leading founding jurists. Tartakovsky is himself the James Wilson Fellow in Constitutional Law at the

About the Author:
Karen Lugo is the Director of the Center for Tenth Amendment Action at the Texas Public Policy Foundation. She was previously Co-Director of the Center for Constitutional Jurisprudence Center at the Claremont Institute in California, a clinical visiting and adjunct professor at Chapman University School of Law, an appointee to the California Advisory Committee to the U.S. Commission on Civil Rights, and president of the Orange County, California, Federalist Society lawyers chapter.
that Witherspoon shared thoughts on morality and philosophy. Jersey, which would later become Princeton. There is evidence his influence through writing and teaching at the College of New He came to America at the behest of Benjamin Rush and extended influence in America. Witherspoon was a signer of the Declaration of Independence and a powerful force in the convention debates. 2 He was one of the Scottish provided a reasoned foundation for civic moral virtue and the principles that undergird liberty. He was one of the intellectual potential that they funded his tuition at King’s College in New York City. When the college closed due to British occupation of the city, Hamilton served George Washington as staff officer—and right hand man—during the Revolutionary War. As early as the middle of the war, Hamilton worried that a weak and disunited government was more of a risk than a powerful, centralized government when it came to the potential usurpation of citizen rights. In addition to Hamilton’s contributions to The Federalist Papers, he wrote a six-part essay series called The Continentalist to fortify the concept of a robust American federal government. Hamilton earned so much credit for influencing New York’s decision to ratify the Constitution that some suggested that New York City be renamed “Hamiltonia.” But Hamilton was not ignorant of the pitfalls of a strong central government, and especially a strong executive; he thought that, should executive power go too far, it could devolve into tyranny. He was especially active in opposing President John Adams’ 1798 Alien and Sedition Acts because of their potential for abuse. In the early 1790s, Hamilton produced three detailed reports on debt, taxes, a national bank, and manufacturing. Historian Gordon S. Wood has noted that Hamilton’s goal in proposing a federal bank was to lend to large commercial enterprises and to

II. Alexander Hamilton

This author concentrates primarily on Wilson’s starring role during the Philadelphia Constitutional Convention, where he was just slightly less significant than James Madison. James Wilson made the first major speech of the Convention, and his 168 substantive speeches at the ratification debates both promoted and elucidated the Constitution. Wilson worked with the Convention’s Committee on Detail to define separated and checked federal powers and the jurisdiction of the federal courts. Wilson’s speeches were recorded and published in an 800-page volume that is full of insights into the structure of the Constitution.

Wilson was a son of the Scottish Enlightenment, which provided a reasoned foundation for civic moral virtue and the principles that undergird liberty. He was one of the Scottish emigrants who arrived with recently distilled philosophy on natural law and lessons on how it might relate to governmental structure. Although he attended university in Scotland intending to become a minister, he later switched his emphasis to law. Tartakovsky mentions the general influence of the Scottish Enlightenment during Wilson’s time in university, and he writes that Wilson especially gained from Thomas Reid’s teaching on “moral sense.” Wilson’s close collaboration with Benjamin Rush, also from the Scottish school, in promoting the Constitution’s adoption, suggests the sustained influence of this distinct worldview on America’s origins. The indispensable influence of the Scottish thinkers is important to recognize, and we can reasonably speculate that John Witherspoon was another source of this influence in America. Witherspoon was a signer of the Declaration of Independence and a powerful force in the convention debates. He came to America at the behest of Benjamin Rush and extended his influence through writing and teaching at the College of New Jersey, which would later become Princeton. There is evidence that Witherspoon shared thoughts on morality and philosophy with John Adams and James Madison. Wilson considered lawful government to be “founded on the law of nature: it must control every political maxim: it must regulate the legislature itself.” Wilson believed that rights exist in nature, and that government exists “to acquire a new security for the possession or the recovery of those rights.” Only this relationship could guarantee the citizen “a natural right to his property, to his character, to liberty, and to safety.” Wilson observed that the conceptual constitutional plan was refined by the state debates, and he thought that the arguments made in the debates should also serve to provide interpretive color: “As the instrument came from [Philadelphia], it was nothing more than a draft plan.” He credited the state conventions with breathing “life and validity” into it.

Unfortunately, historians believe the bulk of Wilson’s personal memoranda was destroyed. It is America’s loss that so few of his personal documents survived.


4 Hadley Arkes, Beyond the Constitution 65 (1990).

5 Id.
provide a platform for enabling international trade. Wood has also surmised that Hamilton's robust military ambitions at home and abroad could be considered Napoleonic.

Tartakovsky cites historian Leonard D. White to laud Hamilton as "one of the great administrators of all time," and he illustrates this with Hamilton's commitment to superintending the Treasury Department's robust growth with an eye to the mission of protecting rights and property. This duality may seem more like a contradiction in light of today's combination of expansive government and capricious infringement of property rights. But Tartakovsky describes Hamilton's mindset as believing the government must have enough power to be able to preserve order and, therefore, liberty.

In a book about how the various characters shaped the Constitution, it would have been useful to read more about Hamilton the lawyer. Supreme Court Justice Joseph Story described Hamilton's legal prowess: "I have heard Samuel Dexter, John Marshall, and Chancellor [Robert R.] Livingston say that Hamilton's reach of thought was so far beyond theirs that by his side they were schoolboys—rush tapers before the sun on noonday." He was considered by colleagues to be "the best trial lawyer of his generation." Inquiring legal minds may want to know more about Hamilton's skill and its bearing on the shape of the Constitution.

One important feature of Hamilton's legal work that is arguably slighted is his use of implied powers to defend the constitutionality of a national bank. Hamilton's famous treatise, Opinion as to the Constitutionality of the Bank of the United States, was published in 1791 to counter Jefferson and others who argued against creating a national bank. Tartakovsky might have discussed how Hamilton's arguments were pivotal in the later McCulloch v. Maryland pleadings.

The author uses the Hamilton v. Jefferson model as a stand-in for the continuing debate over executive power. He points to State of the Union addresses—with their consistent themes of ambitious projects—to demonstrate that the voice of Hamilton still prevails. But challenges to expansive executive authority are now more vigorous than ever. There has been nothing close to political resolution in favor of vast executive power.

It is interesting to read an account of Hamilton's accomplishments told in contrast to Jefferson's record. One wonders why Tartakovsky does not round out the discussion by making it a four-way contest, with John Adams and James Madison completing the square. Adams offered his own case for the balancing of power and institutional checks against ever-feared corruption and personal aggrandizement. Madison and Hamilton were at loggerheads on many pivotal constitutional issues. Gordon Wood observes that Adams and Jefferson were friends, compatriots, and then embittered enemies, but that the one common interest they always shared was hatred for Alexander Hamilton's ambitions and ideas. What was the range of constitutional issues at the heart of these life-long debates?

III. Daniel Webster

Daniel Webster was a force of nature, as those who dared reckon with him learned. He is one of history's greatest legal orators; indeed, Webster may have broken the mold for that class. Webster's rare talents were presaged by early displays of brilliance. His reported memorization of 700 lines of Virgil in one evening was notable even in a time when memorization was a common skill. Webster's entry into the legal arena was as an actor who re-argued cases for spellbound audiences. Josiah Quincy called Webster an "electric force." He applied his formidable talents to over 1,700 cases, 168 of them before the Supreme Court. Many qualified observers lauded his riveting powers of delivery. John Adams said of a Webster oration that it "will be read five hundred years hence with as much rapture as it was heard." Lincoln thought that Webster's Second Reply was the "very best speech that was ever delivered." Tartakovsky compares Webster's ability to tailor his mode of persuasion to his audience to that of Aristotle.

Tartakovsky establishes Webster as a shaper of the Constitution through cases that are staples in most constitutional law casebooks: Charles River Bridge, Ogden v. Saunders, and Dartmouth College. Webster and William Pinckney masterfully applied Alexander Hamilton's treatise on the doctrine of implied powers (discussed above) to prevail in McCulloch v. Maryland, the case that affirmed federal authority to establish a national bank.

Daniel Webster served as a Senator and as Secretary of State under President William Henry Harrison. He also ran for president, but, Tartakovsky explains, he was "fitted to oppose and not to direct," and he was unsuccessful. A useful memorandum from Daniel Webster to the 21st century may be his known aversion to ad hominem attacks, as expressed in his instruction to his son: "I war with principles, and not with men."

IV. Ida B. Wells-Barnette

Born into the Reconstruction Era, this African-American woman was a profile in resilience and tenacity. As an indefatigable civil rights activist and investigative journalist, she was a forceful agent for reform. She fought for criminal defense protections for black defendants and, often to her own peril, exposed lynching practices. She was a stalwart suffragette, and she even nursed her baby while on speaking circuit. Her pursuit of justice—for women and for blacks who were not experiencing promised civil rights protections—would not be denied. This section will be a revelation to many, and The Lives promotes Wells to her rightful historical rank.

V. Woodrow Wilson

There are many conservative and originalist critics of Woodrow Wilson. Tartakovsky is not one of them. He seeks to rehabilitate—or at least suggest that readers reconsider—Wilson's reputation. He refers to Wilson's constitutional scholarship from his years as an academic and his professed reverence
for America’s founders as assurances that Wilson intended to uphold constitutional foundations. Tartakovsky admits Wilson’s dismissal of those who “want to consult their grandparents about everything” and his aversion to going “back to the annals of those sessions of Congress to find out what to do,” but he argues that Wilson’s stated and observable activism were motivated by his regard for Edmund Burke’s teachings. The author believes that when Wilson called Burke the “authentic voice of the best political thought of the English race,” he was indicating a mentorship that comprehensively influenced his actions. But he accepts Wilson’s professions of admiration for Burkean philosophy too readily, rather than probing his life and words for evidence that he meant what he said or that he properly understood his supposed mentor’s teaching.

Woodrow Wilson’s tenure as President evinced overarching fidelity to the evolutionary spirit of Progressivism, a political philosophy that challenged political structures on the basis of social prerogatives. Tartakovsky points out that virtually everyone in federal politics at the time ran on Progressive themes; Theodore Roosevelt had so popularized strains of the movement that no politician could avoid its appeal. And to Tartakovsky, this widespread embrace of Progressivism did not represent a departure from constitutional traditions; rather, he sees the Progressives as bent on recovering the Constitution through updated interpretation.

Wilson argued that the founders’ “Newtonian” vision of government—ruled by unalterable orbits and gravitation based on checks, balances, and branches—was outdated and needed to be replaced by a political construction that was “Darwinian in structure and practice.”10 His desire for such a shift reveals his foundational orientation as untethered from founding principles. Yet Tartakovsky says that this and other similar expressions did not necessarily mean that Wilson wanted a pliable Constitution because Darwin’s theories had not yet been popularized and Wilson must have been thinking in terms of a Burkean approach to adjusting government by slow modification. This Burkean interpretation cannot explain Wilson’s clear derision for vital American precepts. He maintained that “a lot of nonsense has been talked about constitutional departures he did not intend—he fails to square credit—apparently by ascribing to Wilson too much responsibility for constitutional departures he did not intend—he fails to square this with Wilson’s long record of dismissing the very principles that undergird both the Declaration of Independence and the Constitution. Tartakovsky points to the major institutions that Wilson installed—and that have only increased in power and scope—and argues that the lack of significant pushback against the Wilsonian legacy somehow legitimates it. Wilson indeed had a lasting impact on the country, the law, the world, and history. This impact goes far beyond his establishment of the Clayton Antitrust Act, the Federal Trade Commission, and the Federal Reserve Act. And it is impossible to ignore Wilson’s core faith in central planning, which was manifest in his idealistic pursuits including the Treaty of Versailles, the Fourteen Points, and the League of Nations. But a better explanation for the lack of systemic challenge to Wilson’s policies is that any would-be opponent of these initiatives and institutions knows that attempts to reform them have proven futile.

Wilson showed little regard for one of our most cherished constitutional rights: the First Amendment’s right to freedom of speech. For example, he wholeheartedly embraced the Espionage Act of 1917. The measure, occasioned by WWI-era German espionage, made it a crime “to willfully cause or attempt to cause insubordination, disloyalty, mutiny, or refusal of duty, in the military or naval forces of the United States,” or to “willfully obstruct the recruiting or enlistment service of the United States.”12 Subsequently, Wilson supported the Sedition Act, which forbade spoken or printed criticism of the U.S. government, the Constitution, or the flag. Hundreds were jailed during the closing months of World War I for expressing thoughts that officials claimed could aid the Central Powers. Wilson’s administration also blocked some 75 socialist and other periodicals from delivery by mail. Libraries could not offer German-language books, and German-language newspapers were silenced. Eugene Debs, a socialist leader and five-time presidential candidate, spoke disapprovingly of government prosecutions under the Espionage Act and was one of many incarcerated. Debs’ sentence was commuted in 1921 after the repeal of the Sedition Act in 1920.

Wilson most infamously reinstated segregation in government agencies during his presidency. Wilson reportedly told black activist William Monroe Trotter that “[s]egregation is not a humiliation but a benefit, and ought to be so regarded by you gentlemen.”13 Another shameful episode in his political career was when Wilson, as governor of New Jersey, signed a law providing for forced sterilization of “undesirables.”

Finally, Tartakovsky does not elaborate on Wilson’s controversial appointment of Supreme Court Justice Louis Brandeis, beyond broadly ascribing to Brandeis a judicial philosophy of antitrust and separation of business and government. Brandeis, an early social justice activist, was certainly a Constitution-shaping force.

VI. Stephen Field

Tartakovsky does his best work when he provides texture to a life while showing how events and actions in that life shaped the Constitution. That is exactly what he does—to delightful effect—in his vivid account of the life of Justice Stephen Field.

Field grew up in Connecticut and Massachusetts, where his father was a Puritan preacher. While his upbringin...
foreshadow his adult life, it may have been the source for his belief that all possible influence for good should be brought to bear upon the destiny of a state.

When he arrived in San Francisco in 1945, Field observed that a functioning government was needed for the urgent purpose of recording deeds. Field would be part of the solution as the only lawyer northwest of the Yuba River. His first case was based in the law of Mexico. Another arose from claim jumper disputes, and the trial was located in a saloon, which—not unpredictably—resulted in a mass drawing of revolvers. He judged thieves more harshly than murderers in an early expression of something like Rudy Giuliani’s “broken windows” theory, where lesser crimes are corrected in pursuit of order; Field believed that the whole system could fall if horses or purses could vanish without “prompt justice.” Field had a reputation for combining English common law with practical frontier policy. This merging of principle and pragmatism helped to counteract the “might makes right” impulse so common on the frontier, which made settlers fear that the rules of plunder would prevail if legal order was not maintained.

Field rose to sit on the California Supreme Court until President Abraham Lincoln appointed him as the first westerner, and the first Democrat, to the United States Supreme Court in 1863. Tartakovsky marshals a compendium of cases to show Field’s fierce defense of railroads and corporations. He wrote the opinions in Cummings v. Missouri and Ex Parte Garland, using the Declaration of Independence’s phrase “pursuit of happiness” to strike down legislation that restricted property rights. Field believed that “protection of property and persons cannot be separated.”

Field’s famous dissent in The Slaughterhouse Cases, which gutted the Privileges or Immunities Clause, was based in what he called the “right to labor.” He later channeled Adam Smith to compose his clarion defense of economic freedom in the Butchers Union case, which he said stood for “the right to pursue any lawful business or vocation, in any manner not inconsistent with the equal rights of others, which may increase their prosperity or develop their faculties.”

Tartakovsky likens Field’s jurisprudence in his collected dissents to that of Justice Clarence Thomas. Both jurists set up markers and lay cornerstones with “missionary zeal”—despite being a dissenting minority—so that future jurists may build upon their ideas.

Field earned a reputation as “protector of the Chinese” during his California tenure, but Tartakovsky is very critical of Field’s lack of interest in defending the rights of freed black slaves. This criticism is fair, but the close scrutiny applied to Justice Field is surprising when Tartakovsky attempts to contextualize President Wilson’s overt racism.

VII. Robert H. Jackson

Justice Robert H. Jackson was called the “greatest lawyer of the greatest generation,” but he never earned a law degree, and he never achieved his cherished pinnacle: appointment as Chief Justice of the Supreme Court.

As a member of President Franklin Roosevelt’s administration, Jackson was an early cheerleader for the president’s economic experimentation. He vigorously advocated for the constitutionality of New Deal initiatives as Solicitor General and Attorney General. As Attorney General, Jackson defended wartime price controls and internal surveillance as implemented by an energetic wartime executive. And, ironically in light of his later take on this issue in the Steel Seizure Case, Jackson presented a compelling national security brief arguing that the president’s duty to prevent plane construction from being paralyzed overcame the rights of workers to strike at a production facility. When FDR responded to what he saw as the Supreme Court’s intransigence by threatening to pack it with his own nominees, Jackson—still Attorney General—wrote the definitive defense of court-packing in his book, A Struggle for Judicial Supremacy. But the Court realigned on its own when Justice Owen Roberts pivoted to support New Deal legislation before the threatened court-packing happened.

Justice Jackson was appointed to the Supreme Court six months before Pearl Harbor was attacked. Amid reports of Japanese sabotage, the shelling of oil fields near Santa Barbara by a Japanese submarine, and western states refusing admission to Japanese migrants, the Supreme Court voted 9-0 to uphold a Japanese curfew in the Hirabayashi case. But the next term, Jackson dissented in Korematsu, the infamous Japanese internment case. He argued that, although the Court was not in a position to evaluate claims of military necessity, it was nevertheless unconstitutional for the government to hold persons of Japanese origin in camps. He famously wrote:

But once a judicial opinion rationalizes . . . the Constitution to show that the Constitution sanctions such an order, the Court for all time has validated the principle of racial discrimination in criminal procedure and of transplanting American citizens. The principle then lies about like a loaded weapon, ready for the hand of any authority that can bring forward a plausible claim of an urgent need.14

In his most famous Supreme Court opinion, his concurrence in Youngstown Sheet & Tube Co. v. Sawyer (The Steel Seizure Case), Jackson dismissed President Harry Truman’s claim of emergency to justify seizing steel mills to avoid a worker strike during the Korean War. This came as a surprise to many because the government had based its arguments on Jackson’s own brief that argued for executive prerogatives in a similar case involving airplane production during World War II. Jackson’s concurrence outlined a three-level test of presidential authority based on congressional action which has figured prominently in pivotal federal appellate cases most years since 1952.

Chief Justice William Rehnquist, who clerked for Justice Jackson during the Youngstown proceedings, provides clarifying historical context. In his book, The Supreme Court, he wrote that Truman avoided congressional war authorization for the Korean conflict by calling the engagement a police action based in UN prerogatives. Rehnquist also wrote of the weak enthusiasm for the Korean engagement, as it arose less than five years after WWII hostilities concluded. He posited that this was important background for Supreme Court’s new reticence on war powers at the time of Youngstown.

---

Jackson wrote the majority opinion in *West Virginia Board of Education v. Barnette*, upholding Jehovah’s Witness students' right to not participate in the flag salute. Jackson came to the Court after the *Gobitis* decision that allowed school administrators to make flag salutes mandatory even for religious dissenters, and he was one of two new Justices needed to overrule that case. Jackson famously wrote for the six-justice majority “that no official can prescribe what shall be orthodoxy in politics, nationalism, religion, or other matters of opinion . . . .”¹⁵

Justice Jackson took leave from the Court for a year to assume the role of chief prosecutor of the Nuremburg trials. There were controversies over ex post facto lawmaking and how to define war crimes that attended Jackson’s role in the trials. Tartakovsky writes of Jackson’s even temper and political finesse, and he credits Jackson with achieving agreement between the four key Allied nations to merge their different systems of law. He reports that Jackson worked by candlelight behind closed shades at night for concern over snipers. Jackson’s opening and closing statements were said to have ranked with the great state papers of American history.

While Tartakovsky provides fascinating detail about Jackson as a person and judge, he does not discuss one of Jackson’s most constitutionally influential and controversial opinions: *Wickard v. Filburn*. That case drastically expanded Congress’ power under the Commerce Clause. One would expect a mention of this case in a discussion of how the Constitution was shaped.

VIII. Antonin Scalia

Tartakovsky calls Justice Antonin Scalia a “button pusher” and presents as Exhibit A his first dissent as a judge on the D.C. Circuit Court of Appeals, in which he upbraided senior members of the court for being “perverse” and for promulgating reasoning that was “harmful to the national interest.” Scalia’s efforts on the Supreme Court yielded great constitutional dividends. He was “on a mission” to warn Americans that the Supreme Court was slowly expropriating democratic powers, and the author lists an array of social and cultural issues where Scalia saw judicial usurpations of the legislative process. *The Lives* showcases Scalia’s provocative and erudite opinions, books, and public statements to demonstrate the profound impact that the Justice had in his long battle against the “Living Constitution.”

IX. The Finale

Tartakovsky concludes his book on the same hopeful note that echoes throughout his narrative. He suggests that the salutary role of culture will rescue the Constitution when it most needs resuscitation. He points to the social movement that paved the way for legalization of gay marriage as a seminal example of how a trending cause can advance from cultural movement to protected constitutional right. Tartakovsky disregards the legion of legal scholars, some appealed to in the pages of his book, who would argue that this dignity-based license is not even implicitly found in the Constitution.

Furthermore, Tartakovsky fails to reckon with the reality that the Supreme Court’s rulings in cases like *Obergefell* remove controversial issues from the voters and their representatives; the very usurpation of democracy he decries when praising constitutional thinkers like Justice Scalia. Constitutional revolutions are certainly significant, but many would argue that such creative applications of the Constitution chip away at its legitimacy rather than restore it.

¹⁵ 319 U.S. 624 (1943).
A New Foundation for Property Rights?:
A Helpful but Flawed Contribution
by Ilya Somin

A Review of:
Property and Human Flourishing, by Gregory S. Alexander

Note from the Editor:
This review discusses a new book about the foundations of property rights.
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.

Gregory Alexander’s new book Property and Human Flourishing is a major attempt to remake the theory of property rights on new foundations. Instead of justifying property law on the basis of natural rights or utilitarian welfare-maximization, Alexander seeks to ground it on a theory of “human flourishing.” After outlining the foundations of the theory, he then applies it to such varied issues as reparations for historic injustices, the extent to which property owners have the right to destroy their holdings, historic preservation laws, the use of eminent domain, and the alleviation of low-income housing shortages. The book makes many interesting points and is clearly a significant contribution to property theory. But it also has some notable drawbacks.

Most traditional theories of property rights justify them on the basis of natural rights, utilitarian consequentialism, or—occasionally—a combination of the two. Alexander contends that such theories are inadequate, and indeed that no one-dimensional “monist” theory can come close to fully accounting for the value and limits of private property. Instead, he proposes that property be analyzed under a “pluralist” framework intended to promote human flourishing. What does human flourishing consist of? Alexander identifies four key elements: “life, understood to include certain subsidiary values such as health; freedom, understood as including the freedom to make deliberate choices among alternative life horizons; practical reasoning; and sociability” (9).

Alexander’s human flourishing theory is both consequentialist and (at least potentially) paternalistic. It is consequentialist in the sense that it seeks to produce empirical results rather than relying on moral principles whose validity is independent of the consequences of specific policies. It is potentially paternalistic because Alexander argues that people should be provided with these four prerequisites of human flourishing even (at least in many cases) where they may not value them or may prefer to trade them off for other goals. Human flourishing, in his view, is “objectively” valuable, not merely worthwhile only in so far as people want it (e.g. 26-28).

In my view, Alexander is right to argue that an adequate theory of property rights should be “pluralistic,” as he puts it, and that property law should not be based on any one single value to the exclusion of others. But I fear that his own theory is not pluralistic enough. Among other things, he does not sufficiently explain why his version of human flourishing should be privileged over other considerations, and especially not why it should be imposed even on many people who may be willing to cut back on aspects of Alexandrian flourishing in order to pursue other goals. For example, a loner may not want or need as much “sociability” as Alexander posits to be necessary. Some may prefer to exercise their freedom by making most important decisions by intuition, rather than deliberation. And so on. One of the main advantages of strong property rights is the opportunity they give owners (and often others) to pursue values that may not be respected or even understood by majority public opinion. The same goes for values that may not fit even the best formulations of human flourishing.

Alexander’s theory of flourishing is also sometimes difficult to apply. He recognizes that its components may sometimes conflict with each other, thereby necessitating tradeoffs. But he also contends that such tradeoffs are feasible, despite the fact that some of the values integral to the theory are “incommensurable”

About the Author:
with each other (28-35). Alexander argues that we can make “rational” choices among “incommensurable” values by engaging in “practical reasoning,” and that judges deciding property cases can do so by finding “the best interpretation of the lawmakers’ vision of justice animating the rule in question . . . and then [seeing] how the competing and incommensurable moral values best fit together to advance that vision” (34). I am skeptical that real-world judges are likely to accomplish such a herculean task well, especially given limited knowledge and the possibility of ideological and other biases.

This last issue highlights a more general shortcoming of some of the analysis in the book. Alexander emphasizes that the human flourishing theory is a consequentialist approach to property law. Whether it can be effectively implemented depends on whether institutions such as courts, legislatures, and bureaucracies can properly apply it under real-world conditions.

But in discussing the application of the theory to various specific issues, Alexander sometimes loses sight of these crucial institutional questions. For example, Chapter 8 includes a thoughtful discussion of the strengths and weaknesses of historic preservation laws (236-48). Alexander argues that such laws can serve important community interests, but also recognizes that they can potentially impede important development projects. He argues that a well-functioning policy can balance the two objectives against each other, and that to do so “the process for making preservation decisions should be as democratic as possible” (248). Unfortunately, Alexander does not consider whether a “democratic” process can really make these sorts of fine-grained decisions well, given extensive evidence of widespread political ignorance and biased thinking among voters. It seems unlikely that a maximally democratic process would actually perform this function well.

Similarly, Chapter 7 critiques the Supreme Court’s controversial decision in Kelo v. City of New London (2005), which ruled that the Fifth Amendment’s requirement that takings be for a “public use” does not bar the use of eminent domain to condemn private property for transfer to a new private owner in order to promote “economic development.” Alexander argues that the Court was wrong to uphold the taking at issue in the case because it underestimated the value of homes for promoting human flourishing, while overvaluing the “attenuated” economic benefits of the project for which the property was condemned (213-15). Instead of upholding takings for any private enterprise that could potentially benefit the public in some way, the Supreme Court should have weighed the extent to which a taking threatens the “core values” of property, such as the autonomy, security, self-expression, and “responsibility” associated with home ownership (223-29).

I agree with Alexander’s conclusion that Kelo was wrongly decided, and with his more general view that “[t]he public use requirement need not be the anemic doctrine it currently is in the United States” (229). But I wonder whether judges are likely to be able to consistently and rigorously apply the concepts of “core values” and human flourishing advanced by Alexander. Both homes and commercial properties vary greatly in the extent to which they advance the values he references, and homeowners vary in the amount of value they attach to their property. Alexander himself attempts to distinguish between primary residences and secondary ones and argues that the former deserve stronger protection against expropriation than the latter. Given such constraints, a rule-based approach barring or severely limiting all takings for private projects might be preferable to one that attempts to provide special protection to homes or other specific types of property.

An important part of the analytical framework underpinning Alexander’s theory is the idea that many legal limitations on property rights are often justified by obligations of “neighborliness,” a metaphor he adapts from Nancy Rosenblum (70-71). Good neighbors, Alexander suggests, recognize that they have reciprocal obligations to other members of the community and so should accept a variety of constraints (71). Perhaps so. But here too, institutional insight is sometimes lacking.

Robert Ellickson’s classic study of how actual neighbors settle disputes finds that they often prefer to avoid resort to government-enforced legal rules, in favor of informal negotiation and dispute resolution. Settling disputes without involving lawyers and government officials not only reduces litigation and enforcement costs, but is seen as essential to maintaining true neighborliness. As one California rancher told Ellickson, “[b]eing good neighbors means no lawsuits.” That certainly does not prove there should never be formal legal constraints on property rights, or that courts and regulators have no legitimate role to play. But it does suggest that a property theory based on the importance of neighborliness to human flourishing should be wary of imposing extensive legal mandates on property owners.

There is much more to Alexander’s book than can be covered here. Despite the occasional weaknesses of his approach, Property and Human Flourishing includes much valuable material on both the general theory of property rights and a variety of important legal issues. I particularly like the discussion of reparations for groups that have been dispossessed of their land, which has valuable analysis of both why reparations may sometimes be justified and the dangers of trying to return land that has been in other hands for long periods of time (ch. 5). The book is likely to be considered a major achievement of its kind. But it could

---


4. For a discussion of this problem, see id. at ch. 8.

5. Id.


7. Id. at 60. I often highlight this part of Ellickson’s book for my property law students as a way of getting them to consider the potential downsides of litigation.
have achieved still more by incorporating greater sensitivity to the limitations of political and legal institutions.