# TEXAS REVIEW of LAW & POLITICS

VOL. 16, NO. 1 FALL 2011 PAGES 1-210

# SUBSCRIBE TO THE *TEXAS REVIEW OF LAW & POLITICS*

The *Texas Review of Law & Politics* is published twice yearly, fall and spring. To subscribe to the *Texas Review of Law & Politics*, provide the *Review* with your name, billing and mailing addresses.

via telephone: (512) 232-1380
facsimile: (512) 471-6988
e-mail: <subscriptions@trolp.org>
online: <www.trolp.org>
or standard mail:

*Texas Review of Law & Politics*
University of Texas School of Law
727 East Dean Keeton Street
Austin, TX 78705-3299

Annual subscription rate: $30.00 (domestic); $35.00 (international). ISSN #1098-4577.

## REPRINTS

It's not too late to get a copy one of your favorite past articles. See the complete list of the *Review*'s past articles at <www.trolp.org>. Reprint orders should be addressed to:

William S. Hein & Co., Inc.
Hein Building
1285 Main Street
Buffalo, NY 14209

# TEXAS REVIEW of LAW & POLITICS

VOL. 16, NO. 1 FALL 2011 PAGES 1–210


ARTICLES



NOTE

## BOARD OF ADVISORS

---




The *Texas Review of Law & Politics* is a conservative law review. Its mission is to be the prime forum for the discussion and debate of contemporary social issues such as crime, federalism, affirmative action, constitutional history, and religious liberties. The *Review* publishes thoughtful and intellectually rigorous conservative articles—articles that traditional law reviews often fail to publish—that can serve as blueprints for constructive legal reform.

The *Texas Review of Law & Politics* is published two times a year—fall and spring. Annual subscription rates: $30.00 (domestic); $35.00 (international). To subscribe, access the *Review*'s World Wide Web site at <www.trolp.org>, contact the staff at (512) 232-1380, or fax a request to (512) 471-6988. Back issues are available from: William S. Hein & Co., Inc., Hein Building, 1285 Main Street, Buffalo, NY 14209, USA; (800) 457-1986. ISSN #1098-4577.

The *Review* welcomes the submission of articles, book reviews, and notes via standard or electronic mail. All submissions preferably should be submitted in Microsoft Word 2003 format. When submitting a manuscript through standard mail, please include a copy on a compact disk, along with a hard copy, addressed to: *Texas Review of Law & Politics*, 727 East Dean Keeton Street, Austin, Texas, 78705-3299. When submitting a manuscript electronically, please attach the manuscript to an e-mail message addressed to <submissions@trolp.org>.

# TEXAS REVIEW of LAW & POLITICS

VOL. 16, NO. 1 FALL 2011 PAGES 1–210



Founded by Adam B. Ross & Dennis W. Donley Jr. in MCMXCVI

# SUBMISSIONS

The *Texas Review of Law & Politics* welcomes the submission of articles, book reviews, and notes via standard or electronic mail. All submissions should be single-spaced, printed on one side only, with footnotes rather than endnotes. An electronic copy of any accepted manuscript is required.

Citations should conform to the *Texas Rules of Form* (12th ed. 2010) and *The Bluebook: A Uniform System of Citation* (19th ed. 2010). Except when content suggests otherwise, the *Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (12th ed. 2011), *The Chicago Manual of Style* (16th ed. 2010), and *Bryan A. Garner, A Dictionary of Modern Legal Usage* (2d ed. 2001).

Manuscripts submitted electronically should be sent to: <submissions@trolp.org>. Manuscripts submitted via standard mail should be addressed to:

*Texas Review of Law & Politics*
The University of Texas School of Law
727 East Dean Keeton Street
Austin, Texas 78705-3299

# ADVERTISING

The *Texas Review of Law & Politics* is now accepting advertising. We offer the perfect medium to reach a conservative audience interested in contemporary social issues, including crime, federalism, affirmative action, constitutional history, and religious liberties. For more information, please call our offices at (512) 232-1380, fax us at (512) 471-6988, e-mail us at <editors@trolp.org>, or write our Managing Editor at the above address.

## PREFACE

For nearly two years, conservatives have challenged the constitutionality of the individual mandate provision of the Patient Protection and Affordable Care Act, colloquially known as Obamacare. As we go to print, the Supreme Court recently granted certiorari in several Obamacare cases and made the unusual decision to extend the typical one-hour limit for oral argument to an extraordinary five-and-a-half hours. Of particular note is that the Court granted ninety minutes solely for the issue of severability, which involves how much of Obamacare the Court must strike down if it determines any portion of the law unconstitutional.

Kenneth A. Klukowski's timely article, *Severability Doctrine: How Much of a Statute Should Federal Courts Invalidate?*, provides a valuable and comprehensive analysis of this doctrine. Mr. Klukowski carefully traces the evolution of severability doctrine through 135 years of Supreme Court precedent, culminating with the modern two-step test from *Free Enterprise Fund v. Public Co. Accounting Oversight Board* that first examines functionality and then determines whether a shortened statute would still fulfill congressional intent consistent with the legislative bargain. Mr. Klukowski also emphasizes the proper use of severability as a doctrine of judicial restraint and helpful tool to preserve democratic accountability. We hope that Mr. Klukowski's article will influence the Court to strike down most, if not all, of Obamacare.

In another article relevant to Obamacare, Mario Loyola's *Trojan Horse: Federal Manipulation of State Governments and the Supreme Court's Emerging Doctrine of Federalism* describes how states have recently faced an unprecedented degree of federal intrusion into their traditional sphere of responsibility. Mr. Loyola argues that conditional federal grants (e.g., the Medicaid-expansion provisions in Obamacare) and conditional preemption (e.g., EPA's requirement that states alter their air quality programs to account for greenhouse gas emissions or face preemption) threaten the Constitution's framework of dual sovereignty and that only the federal judiciary can provide adequate checks against such coercion.

We are currently in the midst of the first presidential campaign since the Supreme Court's landmark decision in *Citizens United v. Federal Election Commission.* In our next article, *Protecting Speech from the Heart: How Citizens United Strikes Down Political Speech Restrictions on Churches and Charities*, Paul Weitzel discusses how *Citizens United* abrogates the restrictions on political speech by churches and charities and analyzes the

consequences of allowing these groups to enter the political arena. Mr. Weitzel argues that tax deductible charities have a constitutional right to speak about politics.

The past ten years have fostered a close examination of our intelligence capabilities to address failures highlighted by the attacks on September 11, 2011. We conclude the fall issue with *The Necessity of Federal Intelligence Sharing with Sub-Federal Agencies*, a Note by Jason B. Jones, one of the *Review's* articles editors. Mr. Jones supports the inclusion of state and local agencies in intelligence operations and provides recommendations to improve U.S. intelligence capabilities.

We hope you enjoy reading these articles and that they will promote constructive dialog and legal reform. I would like to thank the entire staff for their enthusiasm, hard work, and dedication to the *Review*, as well as Adam Ross, Brantley Starr, and Scott Keller for their continued advice and support.

Shauneen M. Garrahan
*Editor in Chief*

Austin, Texas
December 2011

# SEVERABILITY DOCTRINE: HOW MUCH OF A STATUTE SHOULD FEDERAL COURTS INVALIDATE?

KENNETH A. KLUKOWSKI*



* Research Fellow, Liberty University School of Law. Fellow and Senior Legal Analyst, American Civil Rights Union. B.B.A. 1998, University of Notre Dame; J.D. 2008, George Mason University. The author would like to thank Christopher M. Gacek, William H. Hurd, Gregory G. Katsas, Amanda J. Klukowski, Nelson Lund, John C. Nagle, Hiram S. Sasser, III, and also the Alliance Defense Fund. Any insights are doubtless the result of their input. Any errors are the author's alone. Much of the research contained herein was presented in briefs authored on behalf of the Family Research Council and Members of the United States Congress in healthcare litigation. This Article is dedicated to Remington Xavier Klukowski.

## I. INTRODUCTION

After decades of being a backburner topic, severability doctrine has become a major issue in the federal courts. In 2010, the Supreme Court significantly revised governing doctrine when striking down part of the politically-charged Sarbanes-Oxley Act.[1] Among other current cases raising severability challenges to various statutes, none is more significant than the constitutional challenges to the Patient Protection and Affordable Care Act (PPACA or ACA).[2] As of the time of this writing, one of the federal courts invalidating a central provision of the ACA also held the provision nonseverable, striking down this massive statute in its entirety.[3] However the Supreme Court ends up ruling on these high-profile cases, the contest is of such public prominence—indeed, the healthcare litigation is nothing short of historic, with profound constitutional implications regardless of the outcome—that henceforth severability's importance may be significantly elevated. The existing doctrine is not quite as clear as it has seemed to many courts and commentators, and Congress is likely to become more conscious of the effects that severability clauses and their absence may have, and may even consider including nonseverability clauses in some statutes.

Each time a court strikes down a statutory provision, it must determine whether to invalidate only the unconstitutional provision, or instead whether to invalidate the statute in its entirety or in substantial part. Severability is the doctrine of determining whether part or all of a statute can survive without the invalid provision. Although courts confront on a routine

1. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3151, 3161–64 (2010).

2. Pub. L. No. 111-148, 124 Stat. 119 (2010), *amended by* Health Care and Education Reconciliation Act of 2010, Pub. L. No. 111-152, 124 Stat. 1029 (2010). This statute is often referred to as "Obamacare."

3. Florida *ex rel.* Bondi v. U.S. Dep't of Health & Human Servs., 780 F. Supp. 2d 1256, 1305–06 (N.D. Fla. 2011), *aff'd in part and rev'd in part*, 648 F.3d 1235, 1328 (11th Cir. 2011) (affirming the judgment regarding two provisions but reversing with regard to severability), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400). This litigation is ongoing as of this writing.

basis the decision of how much of a statute to nullify, the doctrine governing this intimidating exercise of judicial power is frequently misapplied.

Striking down an inherently invalid provision is the floor of judicial action, not the ceiling. The general rule is that a federal court should not invalidate more of a statute than necessary.[4] Yet the judiciary's proper role goes beyond simply identifying and neutralizing unconstitutional provisions. Courts must determine the impact of excising the defective provision on the remaining statute to determine whether a broader remedy is appropriate.

The question of severability is ubiquitous, arising whenever a court invalidates a single provision of a multiprovisional statute.[5] "Many laws are unconstitutional, but few are entirely so."[6] Few suggest that statutes should uniformly be stricken down entirely whenever a single clause is found constitutionally infirm. Such an approach would produce severe consequences as Congress passes increasingly large and complex statutes. Severability doctrine is the system showcasing the federal judiciary's ongoing attempts to grapple with this challenge.

The Supreme Court decided its first severability case in 1876,[7] which quickly evolved into asking if Congress would have enacted the challenged statute had it known the invalid provision at issue would be discarded.[8] The Court shortly thereafter added that provisions are nonseverable when retaining the statute without them would create effects not contemplated or intended by Congress.[9] After a half-century of developing the concepts explored in this Article, the Court declared the first clear severability test in 1932.[10] The Court then revised this test in 1987, where in *Alaska Airlines v. Brock* the Court added, "[t]he more relevant inquiry in evaluating severability is whether the statute will function in a *manner*

4. Alaska Airlines v. Brock, 480 U.S. 678, 684 (1987) (quoting Regan v. Time, Inc., 468 U.S. 641, 652 (1984) (plurality opinion)).

5. *See* John C. Nagle, *Severability*, 72 N.C. L. REV. 203, 204 (1993).

6. Kevin C. Walsh, *Partial Unconstitutionality*, 85 N.Y.U. L. REV. 738, 739 (2010).

7. *See* United States v. Reese, 92 U.S. 214, 221 (1876).

8. *See* Trade-Mark Cases, 100 U.S. 82, 98–99 (1879).

9. Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 565 (1902).

10. *See* Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 234 (1932) ("The unconstitutionality of a part of an act does not necessarily defeat or affect the validity of its remaining provisions. Unless it is evident that the Legislature would not have enacted those provisions which are within its power, independently of that which is not, the invalid part may be dropped if what is left is fully operative as a law.").

consistent with the intent of Congress."[11] Then in 2006 the Supreme Court expounded three principles as an underlying rationale to inform severability inquiries.[12]

Most recently, the Supreme Court synthesized decades of cases to restate severability doctrine in the 2010 case *Free Enterprise Fund v. Public Co. Accounting Oversight Board.*[13] This case should be construed as creating a two-step test combining the previous test with various major severability cases. Under *Free Enterprise* Step One, a reviewing court must determine whether all of the remaining provisions of the statute are still fully functional without the constitutionally infirm provision.[14] If so, a court then asks under *Free Enterprise* Step Two whether Congress would be satisfied with the remaining statute, invoking a century of case law concerning whether Congress would have passed the abridged statute.[15]

The Court did not devote adequate space to explain this framework, although this lack of discussion may be partially due to the fact that the Court was not overruling prior precedent, so earlier cases could be studied at length for additional authority. Even so, it is unhelpful that Chief Justice Roberts did not specifically cite and reaffirm *Alaska Airlines*' seminal test of whether the statute still functions in the *manner* Congress intended, since as this Article demonstrates this holding of *Alaska Airlines* is still good law. But *Free Enterprise*'s synthesis is now the current framework, so a scholarly exploration of modern severability doctrine should prove useful.

It is ironic that *Free Enterprise*'s articulation is hardly new; perhaps the foremost law review article on severability—Robert Stern's 1937 Article in *Harvard Law Review*—framed the severability inquiry in precisely that fashion.[16] It just took the

11. 480 U.S. at 685 (emphasis in original).

12. *See* Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 329–30 (2006) ("First, we try not to nullify more of a legislature's work than is necessary . . . . Second . . . we restrain ourselves from rewriting state law to conform it to constitutional requirements even as we strive to salvage it. Third . . . a court cannot use its remedial powers to circumvent the intent of the legislature.") (internal citations and quotations omitted).

13. 130 S. Ct. 3138 (2010).

14. *Id.* at 3161–62.

15. *Id.*

16. *See* Robert L. Stern, *Separability and Separability Clauses in the Supreme Court*, 51 HARV. L. REV. 76, 76 (1937) (stating that a provision is severable "(1) if the valid provisions or applications are capable of being given legal effect standing alone, and (2)

Supreme Court 73 years to catch up (without citing Stern, incidentally). But even Stern does not deserve too much credit, as part of the *Free Enterprise* Court's formulation merely revived a test the Court had previously invoked as far back as 1894.[17]

Conceptually, severability can become an issue in various types of constitutional challenges.[18] As the term is most commonly used it refers to instances in which one provision in a statute is found invalid.[19] As seen in this Article, although severability *per se* consists of determining which otherwise-valid provisions must be nullified alongside an invalid provision, the underlying concepts apply to related areas of remedial actions where courts are engaged in judicial review. It is an intrinsic element of

---

if the legislature would have intended them to stand with the invalid provisions stricken out").

17. *See.*

> It is familiar law that one section or part of an act may be invalid without affecting the validity of the remaining portion of the statute. Any independent provision may be thus dropped out, if that which is left is fully operative as a law, unless it is evident, from a consideration of all the sections, that the legislature would not have enacted that which is within, independently of that beyond, its power.

Reagan v. Farmers' Loan & Trust Co., 154 U.S. 362, 395 (1894).

18. One author compiled these various types of issues, writing that:

> [S]everability becomes an issue when: (1) a party challenges an entire statute, arguing that if any provision of the statute is unconstitutional and nonseverable, the rest of the statute is ineffective; (2) a party argues that a statutory provision is invalid because it is nonseverable from another, purportedly unconstitutional provision of the statute; (3) a *party* contends that an application of a statutory provision is invalid because it is nonseverable from other, unconstitutional applications of the statute; (4) a party argues that a statute is nonseverable, and therefore, another party's constitutional challenge to a provision of the statute would preclude that party from receiving any relief from other provisions of the statute; and (5) a party challenges a statute as being either constitutionally underinclusive or overinclusive. This list is not exclusive, but it depicts some of the situations in which severability becomes an issue in a case.

Nagle, *supra* note 5, at 208–09 (emphasis added) (internal footnotes omitted).

19. This Article discusses severability in evaluating statutes because it is a doctrine of statutory interpretation. Nevertheless, courts also apply a version of this doctrine to substatutory positive law. Courts have conducted severability analyses of regulations. *See, e.g.*, K Mart Corp., v. Cartier, Inc., 486 U.S. 281, 294–95 (1988); Davis Cnty. Solid Waste Mgmt. & Energy Recovery Special Serv. Dist. v. EPA, 108 F.3d 1454, 1459–60 (D.C. Cir. 1997). Courts have also employed this doctrine in evaluating executive orders. *See, e.g.*, Minnesota v. Mille Lacs Band of Chippewa Indians, 526 U.S. 172, 191–95 (1999); *In re* Reyes, 910 F.2d 611, 613–14 (9th Cir. 1990).

As seen throughout the Article, ascertaining Congress's purpose and intent is central to severability. Thus positive law inferior to statutes—such as regulations or executive orders—are categorically subordinate to congressional will, so the method of applying severability doctrine is substantially different than when examining a statute. Nonetheless, these cases show that courts on occasion find rules and principles used in severability analysis useful in evaluating these Executive Branch enactments.

judicial review that whenever a provision in a statute is found invalid, a court must fashion an appropriate remedy. A statutory provision is not an enactment; it is only one part of an enactment. Few would suggest that a court should invalidate an entire statute every time any aspect of the statute is unconstitutional.

Severability clauses in statutes began appearing late in the 1800s, and became commonplace by 1910.[20] Congress began including severability declarations as saving clauses in response to the judiciary's willingness since the 1870s to regularly invalidate statutes in their entirety due to a single faulty provision.[21] These clauses initially proved quite effective.[22] In 1914, after noting that part of a statute might be invalid, the Court rejected the argument that the entire statute should be struck down by noting Congress's intent expressed in the statute's severability clause, and holding that if the challenged provisions were void, they were nonetheless severable.[23]

Severability is often contested only when there is no severability clause. As the leading authority on severability in the early twentieth century explained, "separability clauses are thus now significant only because of their absence. Like articles of clothing, if they are present little attention is paid to them, but if they are absent they may be missed."[24] Still, the presence or absence of a severability clause is but one factor of the court's inquiry; it is not dispositive.[25]

Yet as explained in this Article, such a clause is significant. Including one creates a presumption of severability, but—contrary to what some scholars argue—without a severability

20. Nagle, *supra* note 5, at 222.

21. *See* Comment on Recent Cases, *Constitutional Law: Partial Invalidity of Statutes: Power of Legislature to Alter General Rules of Construction*, 2 CAL. L. REV. 319, 319–20 (1914) (noting the inclusion of severability clauses in several statutes).

22. *See* Nagle, *supra*, note 5, at 222 n.97 ("The earliest legislative statements that statutory provisions should be construed as being severable were taken at face value by the courts." (citing Ohio Tax Cases, 232 U.S. 576, 594 (1914); Yee Gee v. City & Cnty. of S.F., 235 F. 757, 768–69 (N.D. Cal. 1916); Standard Home Co. v. Davis, 217 F. 904, 916 (E.D. Ark. 1914); State *ex rel* Clarke v. Carter, 56 So. 974, 977 (Ala. 1911); *In re* Opinion of Justices, 123 P. 660, 662 (Colo. 1912) (en banc); Michigan Cent. R.R. v. Murphy, 120 N.W. 1073, 1078 (Mich. 1909); Saari v. Gleason, 148 N.W. 293, 295–96 (Minn. 1914); United N.J. R.R. & Canal Co. v. Parker, 69 A. 239, 245 (N.J. 1908); State v. Clausen, 117 P. 1101, 1114 (Wash. 1911); Borgnis v. Falk Co., 133 N.W. 209, 218 (Wis. 1911)).

23. *Ohio Tax Cases*, 232 U.S. at 594.

24. Stern, *supra* note 16, at 122.

25. United States v. Jackson, 390 U.S. 570, 585 n.27 (1968).

clause there is such a presumption only in the lower courts, not the Supreme Court. Absent an express clause, courts should instead freely search for indicia of congressional intent either for or against severing the invalid provision to salvage part or all of the remaining statute.

The synthesis that emerges from modern severability cases is that an invalid provision cannot be severed if it is a major component of the original legislative bargain Congress embodied in the statute. A provision is such an integral provision if, with an eye to the overall purposes of the statute, the truncated statute no longer serves its general purpose because it cannot function in the manner Congress intended without the unenforceable provision. If so, then a court is to conclude that Congress would likely not have enacted the remaining statute in its resulting condition and invalidate the statute in its entirety, as retaining the residual provisions would essentially rewrite the statute into something Congress did not contemplate. Severability ultimately turns on determining the significance of the invalid provision to the overall statutory scheme Congress intended to create; insignificant or incidental provisions are severable, but central provisions[26]—those of major significance—are not.

In previous years one professor noted that some authorities criticize severability doctrine as too malleable,[27] while others consider it too rigid.[28] Either way, some form of this doctrine is essential.[29] As a normative matter, if conceptualized as a two-step inquiry, after *Free Enterprise* it is possible that the new state of severability doctrine strikes the proper balance for a reliably

26. *See* Lars Noah, Essay, *The Executive Line Item Veto and the Judicial Power to Sever: What's the Difference?*, 56 WASH. & LEE L. REV. 235, 237 (1999) ("In some instances, of course, a statute cannot or should not remain in force after a court has invalidated one of its central provisions . . . .").

27. Mark L. Movsesian, *Severability in Statutes and Contracts*, 30 GA. L. REV. 41, 41 (2005) (citing Eugene D. Cross, Comment, *Legislative Veto Provisions and Severability Analysis: A Reexamination*, 30 ST. LOUIS U. L.J. 537, 550–51 (1986); Steven W. Pelak, Note, *The Severability of Legislative Veto Provisions: An Examination of the Congressional Budget and Impoundment Control Act of 1974*, 17 U. MICH. J.L. REFORM 743, 752–53 (1984); Note, *Severability of Legislative Veto Provisions: A Policy Analysis*, 97 HARV. L. REV. 1182, 1183 (1984)).

28. *Id.* (citing Glenn C. Smith, *From Unnecessary Surgery to Plastic Surgery: A New Approach to Legislative Veto Severability Cases*, 24 HARV. J. ON LEGIS. 397, 477 (1987)).

29. *See* Michael C. Dorf, *Fallback Law*, 107 COLUM. L. REV. 303, 370 (2007) ("[N]o workable system of judicial review could function without a large role for severability.").

predictable test that can be applied harmoniously with other interpretive theories.

It is important for constitutional government that courts have an effective severability doctrine to conduct judicial review in a fashion that does not absolve the political branches of their responsibilities. Severability should never provide political cover for Congress. As the Supreme Court explained in its very first severability case:

> It would certainly be dangerous if the legislature could set a net large enough to catch all possible offenders, and leave it to the courts to [eliminate unconstitutional aspects]. This would, to some extent, substitute the judicial for the legislative department of government. The courts enforce the legislative will when ascertained, if within the constitutional grant of power. Within its legitimate sphere, Congress is supreme, and beyond the control of the courts; but if it steps outside of its constitutional limitations, and attempts that which is beyond its reach, the courts are authorized to, and when called upon in due course of legal proceedings must, annul its encroachments upon the reserved power of the States and the people.
>
> To limit this statute in the manner now asked for would be to make a new law, not to enforce an old one. This is no part of our duty.[30]

An effective severability framework is thus one in which Congress skillfully crafts legislation that elected leaders are willing to stand by. This properly narrows the courts' role in judicial review, not routinely sorting through myriad provisions and striking down multiple sections, effectively rewriting statutes and changing the public policy approved by the people's representatives. Such a framework assigns significant force to a severability clause or its opposite, a nonseverability clause, while giving courts more discretion and latitude when Congress deigns not to include any provision expressing its intent on severing invalid provisions.

Part II of this Article surveys the origin and early development of severability doctrine, discussing its four watershed cases and demonstrating how their underlying concepts relate to other areas of remedying statutory invalidities. Part III explores the three principles underlying severability doctrine unanimously

30. United States v. Reese, 92 U.S. 214, 221 (1876).

embraced by the Supreme Court in 2006: minimalism, preserving statutes, and vindicating congressional intent. Part IV sets forth the modern restatement of the doctrine, the two-step test from *Free Enterprise* that first examines functionality and second determines whether a shortened statute would still fulfill congressional intent consistent with the legislative bargain. It also examines two types of severability (total and partial) and the antecedent challenge of how to define a statutory provision. Part V considers the role of a severability clause in dictating whether a presumption of severability exists for a given statute, and also of overcoming the presumption when it exists. Part V also discusses the applicability of a clause not found in the original enactment, and how the current approach to severability is congruent with other areas of statutory interpretation. Then, Part VI explains the proper use of severability as a doctrine of judicial restraint.

## II. DEVELOPMENT OF SEVERABILITY DOCTRINE

Severability has its roots in the case law of the nineteenth century. It is by necessity judge-made doctrine, as it is a system of statutory interpretation for devising judicial remedies when part of a statute is found constitutionally infirm. As such, it is an unavoidable aspect of the power to "say what the law is."[31] When part of a statute is invalid, what then remains of the law?

Like many legal doctrines, severability has had a long evolution. In part due to the fact that it does not carry any inherent political overtones—it is the same when the challenged statute is a gun-control law or an abortion restriction as it is when it is a petroleum extraction program or an airline regulatory measure—severability doctrine has developed in a logical and coherent fashion. Newer precedents smoothly build on prior precedents, and in those few instances where a later case displaces an earlier one, each has been a well-reasoned and noncontroversial modification.[32] Consequently, a chronological survey of severability case law provides a degree of clear and reliable direction on how and when to excise invalid provisions

31. Marbury v. Madison, 5 U.S. (1 Cranch) 137, 177 (1803).

32. Severability varies in federal court depending on whether the challenged statute is federal versus state, since a federal court exercising diversity jurisdiction applies the severability doctrine of the forum state. Stern, *supra* note 16, at 89–94. Thus severability is substantive—rather than procedural—law under the *Erie* doctrine. *See* Erie R.R. v. Tompkins, 304 U.S. 64, 78–79 (1938).

from statutes, despite the fact that severability is relatively esoteric as an interpretive tool.

### A. *Origins of Severability Doctrine*

The doctrine governing severability in American statutory interpretation finds its roots in the nineteenth century. The first suggestions that a multiprovisional statute could be invalidated either partially or *in toto* were made in 1803 and 1829. Thereafter the Supreme Court began developing a doctrine effectuating this concept in a trio of cases in 1876, 1877, and 1879. A coherent inquiry emerged from these cases, though it would be twenty more years before severability became a fully developed doctrine capable of objective and predictable employment.

Courts seemed to assume *sub silentio* that unconstitutional provisions of statutes could be excised from the whole. "[E]very holding of partial unconstitutionality that does not lead to total invalidation necessarily rests on severability, implicitly if not explicitly."[33] It is unclear whether this was originally considered a general rule versus a categorical rule. It was more than sixty years after the adoption of the Constitution before any court is known to have invalidated an entire statute on account of a single provision, and even then it was a state court, not a federal court.[34]

Prior to that, the first instance of a federal court striking down a statute at all was in the iconic case of *Marbury v. Madison*.[35] The Supreme Court never discussed the issue of severability, instead striking down the provision of the Judiciary Act of 1789 authorizing William Marbury to file an original action in the Supreme Court.[36] Marbury sought a writ of mandamus to compel then-Secretary of State James Madison to deliver Marbury's judicial commission to effectuate John Adams's appointment of Marbury as a justice of the peace in the D.C. local courts.[37] The Court never discussed the possibility of

33. Walsh, *supra* note 6, at 741.
34. *See* Nagle, *supra* note 5, at 212.
35. 5 U.S. (1 Cranch) 137 (1803).
36. *Id.* at 173–76.
37. *See id.* at 167–68 (describing the facts of the case).

invalidating the entire statute.[38] In a modern case such a lack of discussion would not be surprising, since *Marbury* was decided on jurisdictional grounds by holding that the provision authorizing an original suit in the Supreme Court purported to confer jurisdiction where the Constitution denied such jurisdiction to the Court.[39] Subsequent Supreme Court precedent makes clear that jurisdictional issues must be resolved as threshold issues before addressing the merits of a case,[40] and where jurisdiction is lacking a court's power is limited to "announcing the fact and dismissing the cause."[41] A significant portion of *Marbury* is *dicta*,[42] such as Chief Justice John Marshall opining on executive privilege doctrine,[43] the canon against

38. This is fortuitous given that the Judiciary Act was the organizing statute of the entire federal judiciary at that point, so wholesale invalidation could have wrought chaos in the federal governmental system.

39. *Marbury*, 5 U.S. (1 Cranch) at 175–76. As seen in these pages of the opinion, the Judiciary Act of 1789 permitted an aggrieved person in William Marbury's situation to file an initial action with the Supreme Court, petitioning the Court to issue a writ of mandamus which under the facts of this case would have been a mandamus order issued to Secretary James Madison, requiring him to deliver Marbury's judicial commission to him so that Marbury could take his seat on the municipal bench.

40. DaimlerChrysler Corp. v. Cuno, 547 U.S. 332, 341–42 (2006). The reason jurisdiction must be addressed first is because federal courts are courts of limited jurisdiction, and so a court begins with the presumption that a lawsuit brought before it is beyond this limited jurisdiction. Kokkonen v. Guardian Life Ins. Co. of Am., 511 U.S. 375, 377 (1994); St. Paul Mercury Indem. Co. v. Red Cab Co., 303 U.S. 283, 288–89 (1938). It is for this reason that the party bringing an issue before a court bears the burden of establishing that the court has jurisdiction, *Cuno*, 547 U.S. at 342, as the presumption against jurisdiction must be overcome before a suit can proceed.

41. Steel Co. v. Citizens for a Better Env't, 523 U.S. 83, 94 (1998) (quoting *Ex parte* McCardle, 74 U.S. (7 Wall.) 506, 514 (1869)).

42. *See* William W. Van Alstyne, *A Critical Guide to* Marbury v. Madison, 1969 DUKE L.J. 1, 6–8. Despite the fact that the authorities cited above indicate the Court should have confined its discussion to its lack of jurisdiction and the need to dismiss the case (invalidating the challenged statutory provision in the process), Chief Justice Marshall nonetheless opined on various constitutional matters that were not in any way necessary to reaching the conclusion that the Court must throw out William Marbury's suit.

43. *Marbury*, 5 U.S. (1 Cranch) at 144–45 (dictum). This doctrine allows for the executive branch to withhold specific information from Congress, even against congressional subpoenas. Congress is constitutionally entitled to information to correctly perform its legislative and oversight responsibilities. McGrain v. Daugherty, 273 U.S. 135, 175 (1927). However, the Constitution also recognizes that the President must be able to maintain secrecy regarding certain information to properly execute his duties as Commander-in-Chief and head of state. United States v. AT&T, 551 F.2d 384, 391 (D.C. Cir. 1974); *see also In re* Sealed Case, 121 F.3d 729, 570 (D.C. Cir. 1997); Ass'n of Am. Physicians & Surgeons v. Clinton, 997 F.2d 898, 909 (D.C. Cir. 1993). Executive privilege balances the tension between these two constitutional forces. *See* United States v. Nixon, 418 U.S. 683, 706–11 (1974). This doctrine remains an underdeveloped doctrine, due in part to the fact that it almost always arises only in politically-charged contexts, concerning conflicts between political actors that can change as a result of each congressional and presidential election. *See generally* Kenneth A. Klukowski, *Making*

superfluities,[44] and the political question doctrine.[45] Yet Marshall does not discuss how much of the statute needed to be invalidated as a result of the flawed provision that granted William Marbury his right of action, instead implicitly finding the flawed provision severable by striking down that provision while leaving the remainder intact. This implication was

---

*Executive Privilege Work: A Multi-Factor Test in an Age of Czars and Congressional Oversight*, 59 CLEV. ST. L. REV. 31 (2011).

44. *Marbury*, 5 U.S. (1 Cranch) at 174 (dictum). This rule that a law's text should be construed such that every word is given its own distinct meaning if possible may have originated as *dictum* involving constitutional interpretation, but has long since been elevated to a holding for both constitutional provisions and statutory provisions. *See* Hibbs v. Winn, 542 U.S. 88, 101 (2004); United States v. Menasche, 348 U.S. 528, 538–39 (1955). This general rule is also called the canon against surplusage, *see* Begay v. United States, 553 U.S. 137, 153 (2008); Lamie v. United States Tr., 540 U.S. 526, 536 (2004), the canon against superfluity, *see* Microsoft Corp. v. i4i Ltd. P'ship, 131 S. Ct. 2238, 2248 (2011), or the canon of antisuperfluousness, *see* Corley v. United States, 129 S. Ct. 1558, 1566 & n.5 (2009).

Under this canon, a court has a "duty 'to give effect, if possible, to every clause and word of a statute.'" *Menasche*, 348 U.S. at 538–39 (quoting Montclair v. Ramsdell, 107 U.S. 147 (1883)); *accord* Moskal v. United States, 498 U.S. 103, 109 (1990). Even a single word must be given meaningful effect, as a failure to give it any effect, or even to assign an effect devoid of practical import, would render it "insignificant, if not wholly superfluous." Duncan v. Walker, 533 U.S. 167, 174 (2001). Thus, "a statute ought, upon the whole, to be so construed that, if it can be prevented, no clause, sentence, or word shall be superfluous, void, or insignificant." Mkt. Co. v. Hoffman, 101 U.S. 112, 115–16 (1879). Courts are "reluctan[t] to treat statutory terms as surplusage" in any context. Babbit v. Sweet Home Chapter of Cmtys. for a Great Or., 515 U.S. 687, 698 (1995); *see also* Ratzlaf v. United States, 510 U.S. 135, 140 (1994). Courts still routinely employ this canon in everyday adjudication. *See, e.g.*, Mackey v. Lanier Collection Agency & Serv., Inc., 486 U.S. 825, 837 (1988) (cautioning "we are hesitant to adopt an interpretation of a congressional enactment which renders superfluous another portion of that same law"); Jones v. Astrue, 650 F.3d 772, 775 (D.C. Cir. 2011).

It should nonetheless also be noted that this rule is not absolute, in that courts can "reject words 'as surplusage' if 'inadvertently inserted or repugnant to the rest of the statute . . . .'" Chickasaw Nation v. United States, 534 U.S. 84, 94 (2001) (quoting KARL N. LLEWELLYN, THE COMMON LAW TRADITION 525 (1960)). There is another caveat as well. "The rule applies only if verbosity and prolixity can be eliminated by giving the offending passage, or if the remainder of the text, a competing interpretation." Bruesewitz v. Wyeth LLC, 131 S. Ct. 1068, 1078 (2011). This *dictum* from *Marbury* has long since become a "cardinal principle of statutory construction." Williams v. Taylor, 529 U.S. 362, 404 (2000).

45. *Marbury*, 5 U.S. (1 Cranch) at 169–70 (dictum). The Supreme Court has subsequently developed the political question doctrine into a multifactor inquiry. *See*:

> Prominent on the surface of any case held to involve a political question is found a textually demonstrable constitutional commitment of the issue to a coordinate political department; or a lack of judicially discoverable and manageable standards for resolving it; or the impossibility of deciding it without an initial policy determination of a kind for clearly nonjudicial discretion; or the impossibility of a court's undertaking independent resolution without expressing lack of the respect due coordinate branches of government . . . .

Baker v. Carr, 369 U.S. 186, 217 (1962); *but see* Nixon v. United States, 506 U.S. 224, 228 (1993) (suggesting that a textual commitment is the predominant factor in determining whether the political-question doctrine applies).

subsequently noted as significant both by the Supreme Court and by federal trial courts,[46] presaging development of an explicit rule.

The first Supreme Court statement beginning to expressly lay the foundation for severability doctrine came in 1829, also from Chief Justice Marshall.[47] In the conclusion of *Bank of Hamilton v. Lessee of Dudley*, Marshall wrote for the Court:

> If any part of [a statute] be unconstitutional, the provisions of that part may be disregarded while full effect will be given to such as are not repugnant to the constitution of the [United States] . . . . The question of whether any of its provisions be of this description, will properly arise in the suit brought to carry them into effect.[48]

*Bank of Hamilton* pronounced what would subsequently become the general rule of severability, that unconstitutional provisions in statutes can be separated from the remainder while leaving the unoffending provisions to continue in effect. Contemporaneously, various state supreme courts began suggesting the negative corollary of *Bank of Hamilton*, that being the possibility of invalidating otherwise-constitutional statutory provisions because these provisions could not be uncoupled from an unconstitutional provision.[49]

The first known case to formally announce and employ the exception to the general rule was a Massachusetts case from 1854.[50] In it, the state's Supreme Judicial Court restated the general rule "that the same act of legislation may be unconstitutional in some of its provisions, and yet constitutional in others."[51] The court then went on to explain:

---

46. *See* Sturges v. Crowninshield, 17 U.S. (4 Wheat.) 122, 201 (1819) (discussing removing a section of bankruptcy law); United States v. The William, 28 F. Cas. 614, 618 n.1 (D. Mass. 1808) (No. 16,700) (supporting the general doctrine of severability); Glenn v. Humphreys, 10 F. Cas. 471, 472 (C.C.E.D. Pa. 1823) (No. 5,480) (stating that an unconstitutional portion of law does not make the entire law unconstitutional). *Accord* Stern, *supra* note 16, at 79 n.9.

47. Nagle, *supra* note 5, at 212 & n.47.

48. 27 U.S. (2 Pet.) 492, 526 (1829).

49. *See* Clark v. Ellis, 2 Blackf. 8, 10 (Ind. 1826) (stating that an unconstitutionality of a provision does not affect the constitutional provisions of the act); Campbell v. Miss. Union Bank, 7 Miss. (6 Howard) 625, 677 (1842) (holding that an unconstitutional act of a bank's charter does not void the remainder of the charter); Exch. Bank v. Hines, 3 Ohio St. 1, 34 (1853) (stating that if an independent provision that is not essential is unconstitutional, it may be treated as void and the rest of the act is enforceable).

50. Nagle, *supra* note 5, at 211.

51. Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 98 (1854).

> Such act has all the forms of law, and has been passed and sanctioned by the duly constituted legislative department of the government; and if any part is unconstitutional, it is because it is not within the scope of legitimate legislative authority to pass it. Yet other parts of the same act may not be obnoxious to the same objection, and therefore have the full force of law, in the same manner as if these several enactments had been made by different statutes. But this must be taken with this limitation, that the parts, so held respectively constitutional and unconstitutional, must be wholly independent of each other. But if they are so mutually connected with and dependent on each other . . . as to warrant a belief that the legislature intended them as a whole, and that, if all could not be carried into effect, the legislature would not pass the residue independently, and some parts are unconstitutional, all the provisions which are thus dependent, conditional or connected, must fall with them.[52]

This invocation and consideration of legislative intent in *Warren v. Mayor of Charlestown* became the basis of nascent severability doctrine,[53] and states began adopting this rule shortly thereafter.[54] The court then concluded that the nature of the law containing the challenged provision was of such a nature "that if this act be unconstitutional at all it is not in any separate and independent enactments, but in the entire scope and purpose of the act."[55] Holding one challenged provision invalid, the Massachusetts court then struck down the statute in its entirety.[56]

---

52. *Id.* at 98–99.

53. Nagle, *supra* note 5, at 213.

54. *See, e.g.*, Lathrop v. Mills, 19 Cal. 513, 530 (1861) (stating that if constitutional provisions are disconnected from the unconstitutional provisions of the act such that the legislature appears to have indented that the constitutional portions of the act be enforced despite the invalidity of the unconstitutional portions, the constitutional provisions are valid); Campau v. City of Detroit, 14 Mich. 276, 285 (1866) (invalidating an entire statute concerning jury duty due to an unconstitutional jury-composition provision); Gordon v. Cornes, 47 N.Y. 608, 616–17 (1872) (upholding an act regarding school funding because the connection between the constitutional and unconstitutional parts of an act are not so connected as to justify the assumption that the legislature did not intend for the constitutional parts of the act to go into effect without the unconstitutional parts of the act.); State *ex rel.* Huston v. Comm'rs of Perry Cnty., 5 Ohio St. 497, 506–07 (1856) (invalidating a local-government statute due to an unconstitutional provision penalizing a county for the location of its seat); Slauson v. City of Racine, 13 Wis. 398, 403–05 (1861) (invalidating an annexation statute due to an unconstitutional taxation provision).

55. *Warren*, 68 Mass. (2 Gray) at 99–100.

56. *See id.* at 101.

The first instance where the Supreme Court of the United States followed a similar approach in striking down a statute *in toto* due to one invalid part appears to be *United States v. Reese*,[57] marking the first in a trio of cases inaugurating severability case law from the High Court. The Court decided *Reese* in 1876, wherein the Court began by holding unconstitutional two sections of a federal statute making it a crime for an election official to refuse allowing a person who is entitled to vote from casting their ballot.[58] In an opinion written by Chief Justice Waite, the Court invalidated this provision because the Fifteenth Amendment—which this statute purportedly was pursuant to—only concerns disallowing discrimination on account of race or former slave status,[59] not any other reasons for which a person might be disenfranchised. The Court then went on to strike down the whole statute, holding that these two sections could not be separated from the other sections.[60] Thus the whole law fell on account of its two unconstitutional components.[61]

57. 92 U.S. 214 (1876).

58. *Id.* at 218–20.

59. *Id.* at 217–18; *see* U.S. CONST. amend. XV.

60. *Reese*, 92 U.S. at 221.

61. It should be noted that the Court's language here includes echoes of the rule of lenity. Under that rule, ambiguities and uncertainties in a criminal law are construed in favor of the defendant. United States v. Granderson, 511 U.S. 39, 54 (1994); Muscarello v. United States, 524 U.S. 125, 139 (1998); Bifulco v. United States, 447 U.S. 381, 387 (1980). This emanates from an overall principle of America as a free society, wherein the benefit of the doubt goes to individual liberty, and so laws entailing possible deprivation of liberty (i.e., criminal statutes by means of incarceration) are construed narrowly on the margins. This is especially true for federal criminal statutes, because whereas states are governments of general jurisdiction with police power to make laws involving public safety, morality, and social welfare, Gonzales v. Oregon, 546 U.S. 243, 270 (2006) (quoting Medtronic, Inc. v. Lohr, 518 U.S. 470, 475 (1996)), the federal government is a government of enumerated powers only, McCullough v. Maryland, 17 U.S. (4 Wheat.) 316, 405 (1819). This is part of the rationale under which there is no such thing as federal common law crimes; all federal crimes must arise from an express statutory enactment. United States v. Hudson & Goodwin, 11 U.S. (7 Cranch) 32, 33–34 (1812). This principle surfaced as recently as 2011, where the Supreme Court held that there is no federal common law cause of action for states to sue for alleged injuries resulting from purportedly-manmade global warming. *See* Am. Elec. Power v. Connecticut, 131 S. Ct. 2527, 2539–40 (2011).

The statute in question in *Reese* was a criminal statute, and the Court's severability analysis suggests that the nature of the statute may have been factored into its analysis. *See Reese*, 92 U.S. at 221 ("It would certainly be dangerous if the legislature could set a net large enough to catch all possible offenders, and leave it to the courts to step inside and say who could be rightly detained, and who should be set at large."). Nonetheless, nothing in the Court's reasoning suggests it is limited to the criminal law, and thus the Court's import is a general part of severability case law when examining statutes of any nature.

Shortly thereafter the Court reinforced the line drawn by this doctrine by finding one statute severable and another nonseverable. One was an 1877 case involving a financial burden imposed by a town on ships accessing its wharf, in which the Court restated the rule as it then existed and then applied it to the case. "Statutes that are constitutional in part only, will be upheld so far as they are not in conflict with the Constitution, provided the allowed and prohibited parts are severable. We think a severance is possible in this case."[62] The second case was an examination of a trademark law in 1879,[63] in which the Court reached the opposite result, reasoning "[i]f we should, in the case before us, undertake to make by judicial construction a law which Congress did not make, it is quite probable we should do what, if the matter were now before that body, it would be unwilling to do . . . ."[64] Collectively these cases gave birth to severability doctrine in the federal system.

As seen below in Part II.B., it would be almost six decades before the Supreme Court finally developed severability doctrine in sufficient detail through a series of cases to the level of articulating a full-orbed analytical framework. The dearth of early (pre-1876) federal case law concerning severability need not raise concerns of illegitimacy, however, or of incompatibility with fidelity to constitutional principles consonant with the Framers' design (for those who are concerned with such things). Severability becomes more of an issue as legislation becomes increasingly long and complex, as courts must consider whether a particular invalid provision is a discrete proposition that can be cleanly separated, versus an integral aspect of the legislation. First, while federal courts have always parsed legal texts, federal judges have had an increasingly large role in statutory interpretation as federal enactments have proliferated.[65] Second, the more provisions in a statute, the greater the number of permutations of severability analyses. Recent years have seen legislative behemoths such as the Dodd-Frank Wall Street

62. Packet Co. v. Keokuk, 95 U.S. 80, 89 (1877). Of course, at this early stage of severability doctrine, it begged the question as to whether the "prohibited parts" were severable as that aspect of Supreme Court jurisprudence was not yet well-developed.

63. Trade-Mark Cases, 100 U.S. 82, 98–99 (1879).

64. *Id.* at 99.

65. *See* Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 409 (1989).

Reform and Consumer Protection Act,[66] which is over 2,300 pages,[67] and the Patient Protection and Affordable Care Act (PPACA or ACA),[68] which is over 2,700 pages.[69] The greater the number of provisions a statute contains, the greater the number of possible permutations of constitutional challenges for a severability analysis. Since such voluminous leviathans were unknown to the Early Republic[70]—indeed, statutes of any length comprised a much smaller portion of governing law in the eighteenth and nineteenth centuries[71]—there were fewer opportunities to consider whether or how invalidating one part of a statute required invalidating the remainder as well. But since the 1980s, statutory interpretation has been an increasingly important topic for the legal academy.[72]

### B. *Evolution of Severability Doctrine*

The cases cited above formed the foundation of severability doctrine. They became so widely known and well-regarded that by 1881, the Supreme Court said in *Allen v. City of Louisiana*, "It is an elementary principle that the same statute may be in part constitutional and in part unconstitutional, and that *if* the parts are *wholly* independent of each other, that which is constitutional may stand while that which is unconstitutional will

66. Pub. L. 111-203, 124 Stat. 1376 (2010).

67. MORTG. BANKERS ASS'N, SUMMARY OF MORTGAGE RELATED PROVISIONS OF THE DODD-FRANK WALL STREET REFORM AND CONSUMER PROTECTION ACT 11 (2010) http://www.mbaa.org/files/ ResourceCenter/MIRA/MBASummaryofDoddFrank.pdf.

68. Pub. L. No. 111-148, 124 Stat. 119 (2010), *amended by* Health Care and Education Reconciliation Act of 2010, Pub. L. No. 111-152, 124 Stat. 1029 (2010).

69. Florida *ex rel.* Bondi v. U.S. Dep't of Health & Human Servs., 780 F. Supp. 2d 1256, 1300 (N.D. Fla. 2011) *order clarified*, 780 F. Supp. 2d 1307 (N.D. Fla. 2011) *aff'd in part and rev'd in part sub nom.* Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health & Human Servs., 648 F.3d 1235 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400). On appeal the Eleventh Circuit stated that the statute was approximately 975 pages. 648 F.3d at 1241. The cause for the apparent discrepancy is that the district court was referring to the statute in legislative drafting form, in which the bill is printed in a double-spaced format with a broad font. The circuit court, by contrast, was referring to the statute as it appears in the Statutes at Large, which are single-spaced in a more compressed font.

70. Daniel A. Farber & Philip P. Frickney, Symposium, *In the Shadow of the Legislature: The Common Law in the Age of the New Republic*, 89 MICH. L. REV. 875, 875 (1991).

71. Movsesian, *supra* note 27, at 43 (citing Frank P. Grad, *The Ascendancy of Legislation: Legal Problem Solving in Our Time*, 9 DALHOUSIE L.J. 228, 251–52 (1985)). In fact, the size of the United States Code essentially tripled between the years of 1964 and 1988. *Id.* (citing W. David Slawson, *Legislative History and the Need to Bring Statutory Interpretation Under the Rule of Law*, 44 STAN. L. REV. 383, 402 (1992)).

72. William N. Eskridge, Jr. & Philip P. Frickey, *Statutory Interpretation as Practical Reasoning*, 42 STAN. L. REV. 321, 321 (1990).

be rejected."[73] The Court then went on to clarify what standard applies if the challenged statutory provision is not "wholly independent" of the other parts, quoting *Warren* from the Massachusetts Supreme Judicial Court cited at length in Part II.A, *supra*, that when provisions are "so mutually connected" as to give rise to the "belief that the legislature intended them as a whole," then "all the provisions which are thus dependent, conditional, or connected must fall" with the invalid provision.[74] The Supreme Court then went on to declare its first version of the test governing severability, the standard being "whether the unconstitutional provisions are so connected with the general scope of the law as to make it impossible, if they are stricken out, to give effect to what appears to have been the intent of the legislature."[75]

Thus began the Supreme Court's jurisprudence on severability. Looking retrospectively just two decades later, the Court declared the principles of severability "well settled,"[76] though unfortunately for his readers Justice John Marshall Harlan did not bother to cite to a single precedent in the Court's discussion of this issue.[77] Helpfully, though, Justice Harlan's opinion for the Court restated the rule as:

> If different sections of a statute are independent of each other, that which is unconstitutional may be disregarded, and valid sections may stand and be enforced. But if an obnoxious section is of such import that the other sections without it would cause results not contemplated or desired by the legislature, then the entire statute must be held inoperative.[78]

This case of *Connolly v. Union Sewer Pipe Co.* invalidated part of an Illinois commercial trust statute on equal protection grounds.[79]

---

73. 103 U.S. 80, 83–84 (1881) (both emphases added).

74. *Id.* at 84 (quoting Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 99 (1854)).

75. *Id.*

76. Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 565 (1902).

77. *See id.* at 564–65.

78. *Id.* at 565.

79. *Id.* at 558–60. The Court's analysis here is consistent with the much-maligned doctrine of Lochner v. New York, 198 U.S. 45 (1905), which was decided only three years thereafter. To the extent that the Court here applied heightened scrutiny in an equal-protection analysis that neither entailed a suspect or quasi-suspect class nor a fundamental right, the Court's holding in the Connolly case was overruled in *West Coast Hotel* and *Carolene Products*. *See* W. Coast Hotel Co. v. Parrish, 300 U.S. 379 (1937); United States v. Carolene Prods. Co., 304 U.S. 144, 152 (1938). *See also* Ferguson v. Skrupa, 372 U.S. 726, 730–31 (1963) (abandoning the *Lochner*-era doctrine).

The Court then held that the invalid provision could not be separated from the remainder of the statute because it would defeat the legislature's intent of granting agricultural companies' special protection, and thus struck down the entire law.[80]

From that day to the present, there have been several significant cases that have expanded and refined severability doctrine, though the approach already explored of examining each invalid provision through the prism of legislative intent to enact an overall scheme is consistently maintained throughout. Moreover, this doctrinal progression has been contemporaneous with the development of related doctrines concerning courts decreeing remedies short of total invalidation when some aspect or application of a statute is held to violate the Constitution.

### 1. Severability Milestones in Supreme Court Case Law

The Supreme Court's decisions in *Allen* and *Connolly* articulate severability doctrine in its nascent state. Although various subsequent cases have included a severability analysis—some having resulted in the offending provisions being severed and others holding that the invalid provisions cannot be severed—four cases in particular have marked significant milestones in this doctrine's evolution.[81]

These came after a series of cases that would have led one to believe that it was a routine matter for a court to invalidate an entire statute on account of a single constitutionally infirm provision. There were at least eleven cases (such as *Connolly*) subsequent to *Allen* in 1881, but prior to the first of these four landmark cases in 1932, in which the Supreme Court held unconstitutional provisions nonseverable, and struck down statutes in whole or substantial part.[82] A survey of case law during

80. *Connolly*, 184 U.S. at 565.

81. Other scholars may well disagree with this particular characterization. There are several cases cited elsewhere in this Article that could reasonably be considered major cases worthy of inclusion on this list of milestones. Conversely, one could reasonably believe two of the cases on this list—*Ayotte* and *Free Enterprise*—to be less important than *Champlin* and *Alaska Airlines*. But each of these four represents distinct developments that must be thoroughly considered in any modern severability analysis, and thus these four form the basis of Part II.B.1.

82. *See* Guinn v. United States, 238 U.S. 347, 366 (1915); Butts v. Merchs. & Miners' Transp. Co., 230 U.S. 126, 135 (1913); Weems v. United States, 217 U.S. 349, 381–82 (1910); Employers' Liability Cases, 207 U.S. 463, 501 (1908); Ill. Cent. R.R. Co. v.

this period shows that it was routine for the Court to strike down otherwise-valid provisions on account of their linkage to invalid provisions.[83] This practice became so widespread that the Court in 1929 held "the general rule is that the unobjectionable part of a statute *cannot* be held separable unless it appears that, 'standing alone, legal effect can be given to it and that the Legislature intended the provision to stand, in case others included in the act and held bad should fall.'"[84] Although presumptions concerning severability will be explored in Part V, case law from this era suggests a presumption *against* severability,[85] one that could only be overcome if a showing of contrary intent could be made. Other courts were more explicit, as illustrated by New Jersey's highest court saying:

> In seeking the legislative intent, the presumption is against any mutilation of a statute, and the courts will resort to elimination only where an unconstitutional provision is interjected into a statute otherwise valid, and is so independent and separable that its removal will leave the constitutional features and purposes of the act substantially unaffected by the process.[86]

This was the state of severability doctrine after its first five decades of development.

### a. *Champlin*

The first landmark severability case is *Champlin Refining Co. v. Corp. Commission of Oklahoma*.[87] An oil refining company challenged several provisions of an Oklahoma statute, arguing that these provisions violated the Commerce Clause and the Fourteenth Amendment's Due Process and Equal Protection Clauses.[88] In determining whether one of these provisions could

---

McKendree, 203 U.S. 514, 529 (1906); James v. Bowman, 190 U.S. 127, 140–42 (1903); *Connolly*, 184 U.S. at 564–65; Pollock v. Farmers' Loan & Trust Co., 158 U.S. 601, 636 (1895); Baldwin v. Franks, 120 U.S. 678, 688 (1887); Spraigue v. Thompson, 118 U.S. 90, 94–95 (1886); Poindexter v. Greenhow (*Virginia Coupon Cases*), 114 U.S. 270, 304 (1884).

83. Stern, *supra* note 16, at 107–08 nn. 138–40 (1937); Alfred Hayes, Jr., *Partial Unconstitutionality with Special Reference to the Corporation Tax*, 11 COLUM. L. REV. 120, 141 (1911).

84. Williams v. Standard Oil Co., 278 U.S. 235, 241 (1929) (emphasis added) (quoting Dorchy v. Kansas, 264 U.S. 286, 290 (1924) (dictum)).

85. *See, e.g.*, Utah Power & Light Co. v. Pfost, 286 U.S. 165, 184–85 (1932) (noting that the common law presumption is that the Legislature intends their acts to be enforced in their entirety).

86. Riccio v. Hoboken, 69 N.J. L. 649, 662 (1903).

87. 286 U.S. 210 (1932).

88. *Id.* at 223–24.

be struck down and then separated from the residue of the oil and gas statute at issue, the Supreme Court declared a general rule of severability that continues to be invoked in 2011:

> The unconstitutionality of a part of an act does not necessarily defeat or affect the validity of its remaining provisions. Unless it is evident that the Legislature would not have enacted those provisions which are within its power, independently of that which is not, the invalid part may be dropped if what is left is fully operative as a law.[89]

Observing that section 10 of the statute was a severability clause, the Court reasoned that the clause "discloses an intention to make the Act divisible and creates a presumption that, eliminating invalid parts, the Legislature would have been satisfied with what remained and that the scheme of regulation derivable from the other provisions would have been enacted without [the invalid provision]."[90] A provision of the statute imposing penalties for violating the Act's section that prohibited committing waste was held void for vagueness under the Due Process Clause.[91] However, the Court severed the unconstitutional provision to preserve the remainder of the statute.

In *Champlin*, the Court promulgated two significant rules for severability inquiries. The first is to articulate severability in a single, albeit complex, sentence, by which to determine whether an invalid provision is severable. "Unless it is evident that the Legislature would not have enacted those provisions that are within its power, independently of that which is not, the invalid part may be dropped if what is left is fully operative as a law."[92] Second, *Champlin* represents a repudiation of part of the previous doctrine that there is a presumption against severability when a statute lacks a severability clause.[93]

89. *Id.* at 234–35 (citing Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 565 (1902); Pollock v. Farmers' Loan & Trust Co., 158 U.S. 601, 635 (1895); Reagan v. Farmers' Loan & Trust Co., 154 U.S. 362, 395–96 (1894); Field v. Clark, 143 U.S. 649, 695–96 (1892)).

90. *Id.* at 235 (citing *Pfost*, 286 U.S. at 165; Crowell v. Benson, 285 U.S. 22, 63 (1932); Williams v. Standard Oil Co., 278 U.S 235, 242 (1929)).

91. *Id.* at 243.

92. *Id.* at 234 (citations omitted).

93. *See*:

> While [a severability clause] is but an aid to interpretation and not an inexorable command, it has the effect of reversing the common law presumption, that the legislature intends an act to be effective as an entirety, by putting in its place the opposite presumption of divisibility; and this

b. *Alaska Airlines*

The next major case is *Alaska Airlines, Inc. v. Brock.*[94] This case is often cited as the modern rule on severability doctrine, though as shown later in this Article that statement is no longer completely accurate, insofar as the *Alaska Airlines* rule has subsequently been developed by later cases and now is the heart of the second part in a two-step inquiry. In *Alaska Airlines*, the Court considered whether a legislative-veto provision of the Airline Deregulation Act of 1978[95]—which was undisputedly invalid after *INS v. Chadha*[96]—was severable from the remaining provisions of the Employee Protection Program set up by the other subsections and paragraphs of section 43 of that Act.[97] In a unanimous opinion written by Justice Harry Blackmun, the Court touched upon various severability precedents[98] and attempted to distill a concise rule to control the question of whether a provision is separable.

*Alaska Airlines* made two significant changes to severability doctrine. First, it took what had been a two-part judicial inquiry and collapsed it into a single test involving legislative intent. The previous articulation was the test from *Champlin*: "Unless it is evident that the legislature would not have enacted those provisions which are within its power, independently of that which is not, the invalid part may be dropped if what is left is fully operative as a law."[99] The *Alaska Airlines* Court reformulated the standard by holding that "[t]he more relevant inquiry in evaluating severability is whether the statute will function in a *manner* consistent with the intent of Congress."[100] This requires a

---

> presumption must be overcome by considerations that make evident the inseparability of the provisions or the clear probability that the Legislature would not have been satisfied with the statute unless it had included the invalid part.

*Pfost*, 286 U.S. at 184–85 (internal citations omitted).

94. 480 U.S. 678 (1987).

95. Airline Deregulation Act of 1978, Pub. L. No. 95-504 § 43(f)(3), 92 Stat. 1705, 1752 (1978).

96. *Alaska Airlines*, 480 U.S. at 680–83 (citing INS v. Chadha, 462 U.S. 919 (1983)).

97. *Id.* at 680–82.

98. *Id.* at 684–86 (citing Regan v. Time, Inc., 468 U.S. 641, 652, 653 (1984) (plurality opinion); *Chadha*, 462 U.S. at 931–32; Buckley v. Valeo, 424 U.S. 1, 108 (1976); Tilton v. Richardson, 403 U.S. 672, 684 (1971); United States v. Jackson, 390 U.S. 570, 585 & n.27 (1968); Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 234, 235 (1932); Hill v. Wallace, 259 U.S. 44, 70–72 (1922); El Paso & Ne. Ry. Co. v. Gutierrez, 215 U.S. 87, 96 (1909)).

99. *Champlin*, 286 U.S. at 234.

100. *Alaska Airlines*, 480 U.S. at 685 (emphasis in the original).

court not only to determine whether the statute would function in some sense, but rather whether it can function in a manner consistent with achieving the purposes for which Congress enacted the statute according to the major dynamics or mechanisms Congress intended to accomplish those purposes. Courts must determine how important the invalid provision was as an element of "the original legislative bargain,"[101] as the terms of that bargain were codified in the statute. "The final test . . . is the traditional one: the unconstitutional provision must be severed unless the statute created in its absence is legislation that Congress would not have enacted."[102] (This test seems to be potentially inconsistent with the inquiry of whether the statute can still function in the manner Congress intended, since it is conceivable that Congress could still pass a statute that does not function in the manner originally intended. This Article shows in Part IV.A.2 how these statements can be reconciled.) The Supreme Court thus granted primacy to the objectives underlying the enactment of the statute, that the legitimizing impetus for codifying public policy in a statute was the lodestone for assessing whether a given provision could be removed from a statute without bereaving the enactment of its legitimizing character as the product of a legislative process consistent with the antecedent premises of a democratic republic.

The Court recast the *Champlin* test to ask whether the statute at bar could function in the manner Congress intended absent the challenged provision.[103] The refinement was necessary, as a legislative veto "by its very nature is separate from the operation of the substantive provisions of a statute,"[104] since all a legislative veto provides is that whatever Executive Branch action has taken place—this action being the manifestation of the statute's substantive provisions—the legislature can negate that action.

The *Alaska Airlines* Court accomplished this change by making statutory functionality part of legislative intent. The second development is the rationale by which the congressional-intent test subsumed the workability prong (which was the second prong) of *Champlin*. "Congress could not have intended a constitutionally flawed provision to be severed from the

101. *Id.*
102. *Id.*
103. *Id.* at 684–85.
104. Nagle, *supra* note 5, at 210.

remainder of the statute if the balance of the legislation is incapable of functioning independently."[105] The Court thereby made what had previously been a separate question—the question of functionality—infer congressional intent. If the issue of whether the residue of a statute is "fully operative as law" as required by *Champlin* factors into the premise that Congress could not have intended to create a dysfunctional statute, then the entire severability inquiry ultimately turns on ascertaining legislative intentions.[106]

The second change to severability doctrine is that the Court effectively abrogated previous cases on the effect of the absence of a severability clause, while also setting the bar to be applied when an express clause is present. As seen above, central to *Alaska Airlines* is determining whether the statute can function in the intended manner without the invalid provision, such that the statute may not have been passed without it:

> The inquiry is eased when Congress has explicitly provided for severance by including a severability clause in the statute. This Court has held that the inclusion of such a clause creates a presumption that Congress did not intend the validity of the statute in question to depend on the validity of the constitutionally offensive provision. In such a case, unless there is strong evidence that Congress intended otherwise, the objectionable provision can be excised from the remainder of the statute.[107]

Contrast this standard that controls when a severability clause is present with the bar set when the legislature elects not to include a clause. "In the absence of a severability clause, however, Congress's silence is just that—silence—and does not raise a presumption against severability."[108]

It immediately became clear that *Alaska Airlines* was the new rule. In *New York v. United States*, the Court considered whether a provision of a federal statute concerning the disposal of radioactive waste that commandeered state legislatures in

105. *Alaska Airlines*, 480 U.S. at 684.

106. Ala. Power Co. v. U.S. Dep't of Energy, 307 F.3d 1300,1307 (11th Cir. 2002); New Haven v. United States, 809 F.2d 900, 907 (D.C. Cir. 1987).

107. *Alaska Airlines*, 480 U.S. at 686 (citations omitted).

108. *Id.* (citations omitted). This last statement is consistent with other aspects of statutory interpretation. *See* Kimbrough v. United States, 552 U.S. 85, 103 (2007) ("Drawing meaning from silence is particularly inappropriate" when "Congress has shown that it knows how to [declare its intent] in express terms.").

violation of the Tenth Amendment could be severed from the Act.[109] In her opinion for the Court, Justice Sandra Day O'Connor expressly noted that the statute contained no severability clause and invoked the rule from *Alaska Airlines* that congressional silence on the issue of severability does not raise a presumption against severance.[110]

c. *Ayotte*

The third landmark severability case is *Ayotte v. Planned Parenthood of Northern New England.*[111] The *Ayotte* Court was considering a New Hampshire statute involving parental notification before a minor could obtain an abortion, which the respondents alleged violated their civil rights under 42 U.S.C. § 1983.[112] This statute included an express severability clause.[113]

This opinion written by Justice Sandra Day O'Connor is especially useful for three reasons. First, it is of recent vintage, and therefore takes into account modern doctrinal developments in statutory interpretation and legislative theory. Second, *Ayotte* is helpful because like *Alaska Airlines*, it was a unanimous decision. Thus, this case expresses articulations regarding severability that at least six of the nine Justices sitting on the Court at the time of this writing claim to be willing to support.

The most useful aspect of *Ayotte*, however, is that it is the first detailed articulation of the underlying rationale for severability doctrine in Supreme Court case law. Justice O'Connor's opinion takes a step back from black-letter rules and discusses the judicial policies implicated by severability inquiries. Surveying various precedents, the Court declared the three following principles:

> Three interrelated principles inform our approach to remedies. First, we try not to nullify more of a legislature's work than is necessary . . . . Second, mindful that our constitutional mandate and institutional competence are limited, we restrain ourselves from 'rewriting state law to conform it to constitutional requirements' even as we strive to

109. 505 U.S. 144, 174–77 (1992).
110. *Id.* at 186 (quoting *Alaska Airlines*, 480 U.S. at 686).
111. 546 U.S. 320 (2006).
112. *Id.* at 323–24 (referencing N.H. REV. STAT. ANN. §§ 132:24–28 (Supp. 2004)).
113. *Id.* at 331 (quoting N.H. REV. STAT. ANN. § 132:28 (Supp. 2004)).

> salvage it. . . . Third, the touchstone for any decision about remedy is legislative intent, for a court cannot 'use its remedial powers to circumvent the intent of the legislature.'[114]

These principles are explored in greater detail in Part III, *infra.* Each provides valuable assistance in fashioning remedies in a severability analysis.

This case has been significantly underutilized to date, although again it is an admittedly recent decision. As seen in various severability analyses and multiple tangentially-related cases from 1932 through 1987, courts tended to cite *Champlin* without much discussion.[115] (Prior to *Champlin,* there was no single test and so severability cases tended to cite to multiple cases to identify settled precedent.[116]) Likewise, since 1987 courts cite to *Alaska Airlines* without elaboration.[117] Yet, part of the impetus behind this Article is that severability doctrine is underdeveloped in some aspects. *Ayotte* helps remedy this deficiency by providing a reasoned approach to the purposes of this doctrine, and so the *Ayotte* Court's holding should provide significantly more guidance to lower courts than most—if not more than any one—of the previous cases.

### d. *Free Enterprise*

The fourth landmark severability case is *Free Enterprise Fund v. Public Co. Accounting Oversight Board.*[118] Although legal briefs submitted in court cases contemporaneously with the writing of this Article continue to cite *Alaska Airlines* as the modern rule, it is in fact *Free Enterprise* that is the most recent restatement of severability doctrine.

In *Free Enterprise,* the Court was considering a challenge to the Sarbanes-Oxley Act of 2002,[119] a statute reforming the public accounting industry.[120] At issue was the appointment process and

114. *Id.* at 329–30 (alterations and citations omitted).

115. *See, e.g.,* Planned Parenthood v. Danforth, 428 U.S. 52, 83 (1976) (containing little explanatory discussion).

116. *See, e.g.,* Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 564–65 (1902) (citing various cases).

117. *See, e.g.,* New York v. United States, 505 U.S. 144, 186 (1992) (citing *Alaska Airlines*).

118. 130 S. Ct. 3138 (2010).

119. Pub. L. No. 107-204, 116 Stat. 745 (2002) (codified in scattered sections of 11, 15, 18, 28, 29 U.S.C. (2006)).

120. *Free Enter.,* 130 S. Ct. at 3147.

tenure provisions for members of the Public Company Accounting Oversight Board (PCAOB)—a board that was created by Sarbanes-Oxley and is answerable to the Securities and Exchange Commission (SEC), and possesses regulatory authority over accounting firms.[121] Specifically, in this case the Court considered whether the provision relating to the removal of PCAOB members violates the Appointments Clause.[122] Although the power to remove is incidental to the power to appoint,[123] it has long been accepted since the inception of the New Deal in the early 1930s that for-cause restrictions on a President's ability to remove an appointee are constitutionally valid.[124] The five Commissioners of the SEC are presidential appointees confirmed by the U.S. Senate for five-year terms,[125] and are removable by the President only for cause.[126] And PCAOB members, in turn, were removable only for cause by a majority vote of the SEC.[127] The Court held that this double for-cause insulation of PCAOB members from the President rendered the board members sufficiently independent from presidential control so as to violate the Appointments Clause.[128] Although the Supreme Court has interpreted the Appointments Clause as allowing for some restrictions on removal, the executive power is vested in the President,[129] and thus presidential control over any executive-branch officers cannot be so attenuated that the President cannot maintain effective supervision over their activities.

---

121. *Id.* at 3147–48.

122. *Id.* at 3147.

123. *Id.* at 3161 (citing, *e.g.*, Sampson v. Murray, 415 U.S. 61, 70 n.17 (1974); Myers v. United States, 272 U.S. 52, 119 (1926); *Ex parte* Hennen, 38 U.S. (13 Pet.) 230, 259–60 (1839)).

124. *See* Humphrey's Ex'r v. United States, 295 U.S. 602, 620 (1935) (upholding a for-cause dismissal restriction on the President's power to remove members of the Federal Trade Commission).

125. Securities Exchange Act of 1934, Pub. L. No. 73-291 § 4(a), 48 Stat. 881, 885 (1934) (codified at 15 U.S.C. § 78(d)(a) (2006)).

126. Although the 1934 statute is silent on any removal of Commissioners prior to the expiration of their term, *Free Enter.*, 130 S. Ct. at 3182–84 (Breyer, J., dissenting), the Court regarded such a for-cause caveat implicit in the Securities Exchange Act, *see id.* at 3151–54 (majority opinion) (quoting in part *Humphrey's Ex'r*, 295 U.S. at 620)). Courts have criticized reading such implicit for-cause protections into statutes creating independent executive agencies. *E.g.*, *In re* Aiken Cnty., 645 F.3d 428, 447 & n.7 (D.C. Cir. 2011) (Kavanaugh, J., concurring).

127. *See Free Enter.*, 130 S. Ct. at 3146 (citing 15 U.S.C. § 7211(e)(6) (2006)).

128. *Id.* at 3151.

129. U.S. CONST. art. II, § 1, cl. 1.

Given that this invalid provision was only one clause in a lengthy and complex statute, the Court then considered whether the invalid provision could be severed from the remainder of the Act.[130] The Court held the double for-cause removal provision severable from the remainder of Sarbanes-Oxley.[131] The decision was correct, as even the challengers to Sarbanes-Oxley acknowledged that their challenge to the legality of the Board's appointment and removal provision was "collateral" to any orders or regulations promulgated by the Board to which the plaintiffs might object.[132] In so doing, the Court modified *Alaska Airlines*' severability inquiry.

The Court began by quoting *Ayotte*. "Generally speaking, when confronting a constitutional flaw in a statute, we try to limit the solution to the problem."[133] The Court in *Free Enterprise* then employed what amounts to a two-step inquiry: First, the remainder of the statute must continue to be "fully operative as a law" absent the invalid provisions.[134] If the remainder would be fully operative, the second step is to uphold the truncated statute "[u]nless it is evident that the Legislature would not have enacted those provisions . . . independently of that which is [invalid]."[135]

The Court's methodology in reviewing the Sarbanes-Oxley Act confirms this new two-step approach. After discussing how the functioning of the public board at issue in that case would continue unaffected by invalidating the removal mechanism of board members, the Court concluded that the "Act remains 'fully operative as a law' with these tenure restrictions excised."[136] Only then did the Court consider congressional intent, invoking the language from *New York* and *Alaska Airlines*.[137] Having articulated the framework, the Court then took each step in turn. The Court first held that the remainder of the statute was capable of functioning independently of the

130. *Free Enter.*, 130 S. Ct. at 3161–64.

131. *Id.* at 3161.

132. *Id.* at 3150.

133. *Id.* at 3161 (quoting Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 328 (2006)).

134. *Id.* (quoting New York v. United States, 505 U.S. 144, 186 (1992) (quoting in turn Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 684 (1987))).

135. *Id.* (quoting *New York*, 505 U.S. at 186) (ellipsis and alterations in the original).

136. *Id.*

137. *Id.* (quoting *New York*, 505 U.S. at 186).

invalid provision, and then added that the statutory text and legislative historical context did not indicate Congress would have preferred the entire statute to fail.[138] Thus the Court begins with a textual analysis by inquiring whether the abridged statute would still be fully operative. If operative, then the Court proceeds to the second step of determining legislative intent, an admittedly more difficult inquiry with greater potential for judicial error.

This modifies the *Alaska Airlines* test in two respects. First, whereas *Alaska Airlines* had collapsed *Champlin*'s two-part inquiry into a single test, *Free Enterprise* resegregates the inquiry into a functionality prong and an intent prong. Second, *Free Enterprise* inverts the order of the *Champlin* inquiry. A court now first considers whether the statute is still literally functional. If the statute can still function sans the invalid provision, only then does a court inquire into legislative intent, such as whether the statute can still fulfill Congress's purpose in a manner acceptable to Congress or whether Congress would still have enacted a bill without the invalid provision. This change reorients the judicial inquiry to focus on principles of statutory interpretation and construction when possible, rather than explore hypotheticals of how Congress would react under changed facts.

This new inquiry—and the practical consequences thereof—is explored further in greater detail in Part IV.A, *infra.*

### 2. Related Doctrinal Developments: Overbreadth, Facial, and As-Applied Challenges

One aspect in which activity in the judicial branch differs from the legislative branch is in the interconnectedness of legal principles. Statutes are stand-alone legislative enactments, each of which has self-contained force and only modifies the corpus of legislative and regulatory law, as found in the *United States Code* and the *Code of Federal Regulations*, respectively, insofar as it either adds to or amends specific statutory provisions of the U.S. Code (either explicitly or by implication) or supersedes inconsistent prior regulations. Similar to the concept that court decisions are supposed to be based on the application of neutral principles that should span the full spectrum of legal subject

138. *Id.* at 3162.

matters,[139] so too doctrinal developments have applications in other areas of judicial activity, as courts seek to harmonize rules of law.

Two aspects of severability doctrine suggest an underlying dynamic that points to other doctrines. First, broadly speaking, severability is a remedies question, as severability is triggered when judges deliberate on the appropriate remedy when a statutory provision is found unconstitutional. Second, and more narrowly defined, severability is a question of the breadth of invalidity. When an aspect of a statute is found unconstitutional, how broad should the Court's holding sweep in removing the unconstitutionality? In framing the issue in *Ayotte*, Sandra Day O'Connor began, "If enforcing a statute that regulates access to abortion would be unconstitutional in medical emergencies, what is the appropriate judicial response? We hold that invalidating the statute entirely is not always necessary or justified, for lower courts may be able to render narrower declaratory and injunctive relief."[140] There are two doctrines under which courts consider similar questions, doctrines expressing a form of judicial modesty by not invalidating more of a statute than necessary.

The first involves facial versus as-applied challenges.[141] The seminal early law review Article on severability—Robert Stern's 1937 Article in *Harvard Law Review*—characterized judicial distinctions between facial and as-applied challenges as a form of severability inquiry.[142] Subsequent sources have made similar comparisons.[143] When the validity of a statute is challenged, courts often must determine whether to invalidate the statute under all circumstances versus only for certain types of plaintiffs or under specified circumstances. The former is a facial

139. *Cf.* Robert H. Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1, 1–9, 19, 35 (1971) (arguing that the Supreme Court's constitutional role is justified only if the Court applies principles that are neutrally derived, defined, and applied).

140. Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. at 323.

141. Although these two types of challenges are distinct, it should be noted at the outset of this description that the Court recently reiterated as a caveat that "the distinction between facial and as-applied challenges is not so well defined that it has some automatic effect or that it must always control the pleadings and disposition in every case involving a constitutional challenge." Citizens United v. FEC, 130 S. Ct. 876, 893 (2010).

142. Stern, *supra* note 16, at 79; *accord* Walsh, *supra* note 6, at 739.

143. *See, e.g.*, Gillian E. Metzger, *Facial Challenges and Federalism*, 105 COLUM. L. REV. 873, 885 & n.52 (2005) (comparing "application severability" to "text severability").

challenge and the latter is an as-applied challenge.[144] The Court recognized even in the nineteenth century that a law could be valid in many applications—which in modern jurisprudential parlance is to say that the law is facially valid—while other applications would be unconstitutional. One such case, *Reagan v. Farmers' Loan & Trust Co.*, is especially illustrative here because it first distinguishes facial from as-applied challenges,[145] then goes on to also announce the basic rule of severability that invalid provisions can often be separated from valid ones.[146] Upholding a statute facially does not foreclose future challenges that the statute is invalid as applied to the particular challenger and facts of a future case.[147] Courts invalidate statutes facially "sparingly and only as a last resort."[148] This is because—similar to severability—facial challenges are disfavored relative to as-applied decisions as a policy of judicial restraint.[149] The Court has articulated two different standards to succeed in a facial challenge, one being "that no set of circumstances exists under which the Act would be valid,"[150] the other "that the statute lacks

144. *Compare* United States v. Salerno, 481 U.S. 739, 745 (1987) ("A facial challenge to a legislative Act is, of course, the most difficult challenge to mount successfully, since the challenger must establish that no set of circumstances exists under which the Act would be valid."), *with* Tennessee v. Lane, 541 U.S. 509, 552 n.11 (2004) (referencing upholding the statute "as applied to respondents and declin[ing] to entertain the facial challenge"), *and* United States v. Playboy Entm't Grp., 529 U.S. 803, 834 n.3 (2000) (Scalia, J., dissenting) (noting that an as-applied challenge only applies to particular facts of the case, unlike a facial challenge).

145. 154 U.S. 362, 390 (1894) ("A valid law may be wrongfully administered by officers of the state, and so as to make such administration an illegal burden and exaction upon the individual.").

146. The Court continued:

> It is familiar law that one section or part of an act may be invalid without affecting the validity of the remaining portion of the statute. Any independent provision may be thus dropped out, if that which is left is fully operative as law, unless it is evident, from a consideration of all the sections, that the legislature would not have enacted that which is within, independently of that beyond, its power.

*Id.* at 395.

147. *See* Wis. Right to Life, Inc. v. FEC, 546 U.S. 410, 411–12 (2006) ("In upholding § 203 against a facial challenge, we did not purport to resolve future as-applied challenges.").

148. Nat'l Endowment of the Arts v. Finley, 524 U.S. 569, 580 (1998) (internal quotation marks omitted).

149. *See* FW/PBS, Inc. v. City of Dall., 493 U.S. 215, 223 (1990) (noting that "facial challenges to legislation are generally disfavored").

150. *Salerno*, 481 U.S. at 745; *but see* Michael C. Dorf, *Facial Challenges to State and Federal Statutes*, 46 STAN. L. REV. 235, 294 (1994) (arguing that in practice the Supreme Court's standard for facial versus as-applied challenges is different from what the Court held in *Salerno*).

any plainly legitimate sweep."[151] (Conveniently, "[w]hich standard applies in a typical case is a matter of dispute."[152]) Facially invalidating a statute subsumes all possible claims, necessarily holding that all conceivable as-applied challenges to the statute would succeed.[153]

The second is overbreadth doctrine. This doctrine is distinctive to challenges brought under the Free Speech Clause.[154] Under this doctrine, a law burdening free speech is invalid if in regulating speech that is properly regulable, it also restricts a substantial amount of protected speech.[155] Otherwise analogized to the points of law in the previous paragraph, it is a "type of facial challenge in the First Amendment context under which a law may be overturned as impermissibly overbroad because a substantial number of its applications are unconstitutional."[156] Such an enactment is impermissible because it is overly broad, and is facially invalidated. The Court has explained that overbreadth only applies where First Amendment violations are alleged because of the "chilling effect" that laws penalizing speech can have on those who are uncertain whether their contemplated speech is protected.[157] This is a *sui generis* anomaly to the normal rule of partial invalidation, whereby a court invalidates the Act only as applied to the challengers and reserves facial challenges to those where no valid applications can be conceived.[158]

---

151. United States v. Stevens, 130 S. Ct. 1577, 1587 (2010) (quoting Washington v. Glucksberg, 521 U.S. 702, 740 n.7 (1997) (Stevens, J., concurring in judgments)) (internal quotation marks omitted).

152. *Id.*

153. *See* Richard H. Fallon, Jr., *Commentary: As-Applied and Facial Challenges and Third-Party Standing*, 113 HARV. L. REV. 1321, 1339–40 (2000) (dicussing the binding effect of decisions on facial challenges).

154. *See* Schall v. Martin, 467 U.S. 253, 268 n.18 (1984) (noting that "outside the limited First Amendment context, a criminal statute may not be attacked as overbroad").

155. Broadrick v. Oklahoma, 413 U.S. 601, 615 (1973).

156. Wash. State Grange v. Wash. State Republican Party, 552 U.S. 442, 449 n.6 (2008) (internal citation omitted); *accord Stevens*, 130 S. Ct. at 1587.

157. Richard H. Fallon, Jr., *Making Sense of Overbreadth*, 100 YALE L.J. 853, 855–56 (1991).

158. Interestingly in the context of this Article, however, the Supreme Court made clear that overbreadth does not trump severability. In *Brockett v. Spokane Arcades*, the Court held that an invalid provision of a statute that was severable could be separated from the remainder of the Act, and in doing so spared the residue from an overbreadth challenge. 472 U.S. 491, 506 (1985).

Although not expressly invoked, this approach sounds in the doctrine of constitutional avoidance. *See infra* note 160 and accompanying text. By severing the invalid provision the Court was able to excise the provision that could have invalidated

There are additional doctrines founded upon the same rationale of judicial respect for the political branches, such as the rule requiring courts to interpret a statute to avoid unreasonable results when practicable.[159] Another is the doctrine of constitutional avoidance, that a court should give a narrowing construction to a statute when possible to avoid potential constitutional problems,[160] if the statute is "readily susceptible" to such a construction.[161] Severability is of a piece with these commonly-invoked doctrines.[162] In fact, regarding this last doctrine, Chief Justice Charles Hughes invoked the rationale

---

the statute for violating the First Amendment as overbroad, thereby saving most of the statute.

159. Am. Tobacco Co. v. Patterson, 456 U.S. 63, 71 (1982) ("Statutes should be interpreted to avoid untenable distinctions and unreasonable results whenever possible.").

160. FCC v. Fox TV Stations, Inc., 129 S. Ct. 1800, 1811 (2009); Edward J. DeBartolo Corp. v. Fla. Gulf Coast Bldg. & Constr. Trades Council, 485 U.S. 568, 575 (1988) (citing NLRB v. Catholic Bishop of Chi., 440 U.S. 490, 499–501, 504 (1979)); Blodgett v. Holden, 275 U.S. 142, 148 (1927) (Holmes, J., concurring) (positing that "as between two possible interpretations of a statute, by one of which it would be unconstitutional and by the other valid, [the judiciary's] plain duty is to adopt that which will save the Act"). This doctrine has evolved slightly over the decades, as the early version "requires the court to determine that one possible interpretation of the statute *would* be unconstitutional, while the latter requires only a determination that one possible reading *might* be unconstitutional." Adrian Vermeule, *Saving Constructions*, 85 GEO. L.J. 1945, 1949 (1997) (emphasis in the original). Vermeule cites the following as the shift to modern doctrine:

> [U]nless [the early doctrine] be considered as meaning that our duty is to first decide that a statute is unconstitutional, and then proceed to hold that such ruling was unnecessary because the statute is susceptible of a meaning which causes it not to be repugnant to the Constitution, the rule plainly must mean that where a statute is susceptible of two constructions, by one of which grave and doubtful constitutional questions arise and by the other of which such questions are avoided, our duty is to adopt the latter.

United States *ex rel.* Att'y Gen. v. Del. & Hudson Co., 213 U.S. 366, 408 (1909), *quoted in* Vermeule, *supra* note 160, at 1958.

161. Reno v. ACLU, 521 U.S. 844, 884 (1997); *accord* Machinists v. Street, 367 U.S. 740, 749–50 (1961) ("When the validity of an act of the Congress is drawn in question, and even if a serious doubt of constitutionality is raised, it is a cardinal principle that this Court will first ascertain whether a construction of the statute is fairly possible by which the question may be avoided[,]" quoting Crowell v. Benson, 285 U.S. 22, 62 (1932)) (internal quotation marks omitted). This doctrine is traced to Justice Louis Brandeis. *See* Ashwander v. TVA, 297 U.S. 288, 348 (1936) (Brandeis, J., concurring) (quoting *Crowell*, 285 U.S. at 62).

162. *See* Vermeule, *supra* note 160, at 1945–46, 1952–55 (referring to the scholarly consensus that these two doctrines are related). Vermeule's discussion of severability is actually more focused on as-applied challenges, however, so though the discussion is entirely relevant to the issue at hand, the author's comparison of severability with the doctrine of constitutional avoidance is not fully consonant with the model set forth in this Article. It should also be noted that Vermeule then argues this consensus is incorrect, and that severability should be regarded as inconsistent with constitutional avoidance. *See id.* at 1955–63.

of the doctrine of constitutional avoidance in *Jones & Laughlin Steel* in terms that can be imported as a rationale for severability doctrine as a manifestation of judicial modesty.[163]

### C. *Severability Consistent with Principled Judicial Inquiries*

These concepts of severability are also consistent with the full range of methodologies of judicial action, also referred to as schools of legal interpretation. While some of these methodologies rest on neutral principles regardless of outcome, and others focus increasingly on outcome regardless of antecedent rules, severability is not stereotyped into any one of them. Severability inquiries can be either neutral or teleological.

There are two such neutral inquiries, so called because they take no account (i.e., are "neutral") regarding the results they produce.[164] The first is textualism, which interprets a legal text according to the meaning of the precise words of the text, subordinating legislative history or other extrinsic or derivative material to the unambiguous meaning of its terms.[165] "One

163. NRLB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 30 (1937) ("The cardinal principle of statutory construction is to save and not to destroy. We have repeatedly held that as between two possible interpretations of a statute, by one of which it would be unconstitutional and by the other valid, our plain duty is to adopt that which will save the act."); *see also* Stern, *supra* note 16, at 105 (citing Justice Hughes's argument against invalidating the National Labor Relations Act *in toto*).

164. This is not to say neutral inquiries are blind to ludicrous results. A *reductio ad absurdum* is a *reductio ad absurdum.* Principled legal interpretation need not produce such results. *See Patterson,* 456 U.S. at 71 (holding that interpretations resulting in unreasonable results and untenable distinctions should be avoided). Statutory interpretation is not an intellectual suicide pact. Of course, those focused on achieving particular outcomes can point to what they regard as the objectionable results of a neutral methodology, and hyperbolically refer to such results as irredeemably absurd. So characterizing this form of inquiry requires a measure of reasonable discretion and self-restraint on the part of those disagreeing with various outcomes.

165. John F. Manning, *Textualism and Legislative Intent,* 91 VA. L. REV. 419, 420 (2005). Less constrained than originalism (described *infra* note 167), textualism is not confined to the precise meaning of those words at the moment of their codification, but eschews common indicia of legislative intent such as floor statements by legislators. *See* Conroy v. Aniskoff, 507 U.S. 511, 519 (1993) (Scalia, J., concurring) ("We are governed by laws, not by the intentions of legislators . . . . 'The law as it is passed is the will of the majority of both houses, and the only mode in which that will is spoken is in the act itself . . . .'" (added emphasis omitted) (quoting Aldridge v. Williams, 44 U.S. (3 Howard) 9, 24 (1845))). Textualism is consistent with Justice Holmes's talismanic statement: "We do not inquire into what the legislature meant; we ask only what the statute means." Oliver Wendell Holmes, *The Theory of Legal Interpretation,* 12 HARV. L. REV. 417, 419 (1899).

This is because some degree of speculation is inherent in proclaiming an intent for any organizational body comprised of many individuals. Frank H. Easterbrook, *Text, History, and Structure in Statutory Interpretation,* 17 HARV. J.L. & PUB. POL'Y 61, 68 (1994). However, it should also be noted that there are variations of textualism that permit consulting legislative history to resolve ambiguities if the legal text is anything short of

determines what Congress would have done by examining what it did."[166] The other is originalism, a form of textualism that requires interpreting legal text in accordance with the original meaning of those words.[167]

---

transparent. *See* Elliot M. Davis, Note, *The Newer Textualism: Justice Alito's Statutory Interpretation*, 30 HARV. J.L. & PUB. POL'Y 983, 986–87, 991–93 (2007) (recounting Justice Alito's use of legislative history in statutory interpretation); *see also*, e.g., Zedner v. United States, 547 U.S. 489, 500–02 (2006) (using legislative history to interpret the Speedy Trial Act).

166. Legal Servs. Corp. v. Velazquez, 531 U.S. 533, 560 (2001) (Scalia, J., dissenting).

167. Originalism interprets words as they would have been understood at the moment of their adoption. An originalist inquiry assigns to each term what a person of ordinary intelligence and education who was reasonably aware of current events would believe them to mean. *See* Attorney General Edwin Meese III, Speech Before the D.C. Chapter of the Federalist Society Lawyers Chapter (Nov. 15, 1985), *in* ORIGINALISM: A QUARTER-CENTURY OF DEBATE, at 71–82 (Stephen G. Calabresi ed., 2007). One scholar derisively referred to this as an "archaeological" approach. *See* T. Alexander Aleinokoff, *Updating Statutory Interpretation*, 87 MICH. L. REV. 20, 21 (1988) (citing Charles Curtis, *A Better Theory of Interpretation*, 3 VAND. L. REV. 407, 415 (1950)). Recent Supreme Court Terms have contained several quintessential examples of originalist analysis. *See, e.g.*, Brown v. Entm't Merchs. Ass'n, 131 S. Ct. 2729, 2751–61 (2011) (Thomas, J., dissenting) (asserting that the Constitution should be interpreted as it was understood at its adoption); McDonald v. City of Chi., 130 S. Ct. 3020, 3058–88 (2010) (Thomas, J., concurring in part and concurring in judgment) (arguing that constitutional interpretation should be based on original understanding even if that requires overturning precedential case law).

These two examples suggest a larger point that over the past decade Justice Clarence Thomas has emerged as a consistent originalist—and the only consistent originalist—currently on the Court. To the two cases referenced above—the former concerning the Free Speech Clause and the latter concerning the Second Amendment and the Fourteenth Amendment Privileges or Immunities Clause—Justice Thomas has also issued separate opinions expressing an originalist position on the Commerce Clause, *see* United Haulers Ass'n, Inc. v. Oneida-Herkimer Solid Waste Mgmt. Auth., 550 U.S. 330, 349 (2007) (Thomas, J., concurring in judgment); United States v. Lopez, 514 U.S. 549, 584–602 (1995), and the Takings Clause, *see* Kelo v. City of New London, 545 U.S. 469, 505–523 (2003) (Thomas, J., dissenting). It is worth a comparative study of Thomas's approach to originalism to that of Justice Antonin Scalia, who might to some degree share Thomas's interpretive philosophy but regards stare decisis as protecting non-originalist doctrines that Thomas is unconstrained to overrule. *See, e.g.*, Dep't of Revenue v. Davis, 553 U.S. 328, 359 (2008) (Scalia, J., concurring in part) ("I will apply our negative Commerce Clause doctrine only when stare decisis compels me to do so."). This material is noted because it is quite possible that these debates of textualism versus originalism and the proper role of stare decisis when considering longstanding doctrines will be increasingly prominent in legal debate.

On a related note, there is more potential for originalism to impact the jurisprudential development of constitutional provisions for which there are relatively few precedents, as stare decisis presents less of a barrier to interpreting and applying those provisions. The Second Amendment presents such an opportunity. *See, e.g.*, Kenneth A. Klukowski, *Citizen Gun Rights: Incorporating the Second Amendment Through the Privileges or Immunities Clause*, 39 N.M. L. REV. 195, 234–52 (2009) (applying an originalist understanding of the Fourteenth Amendment to the Second Amendment right to keep and bear arms); Nelson Lund, *The Second Amendment, Heller, and Originalist Jurisprudence*, 56 UCLA L. REV. 1343, 1344–48, 1368–76 (2009) (discussing the original meaning of the right to keep and bear arms).

Under either of these, a court is still giving effect to the remaining terms of the statute in accordance with their meaning. If it is a *Free Enterprise* Step One inquiry (explained *infra* in Part IV.A), the examination is limited exclusively to the words of the statute regardless, and so there is no occasion to consult extraneous sources of contested value. If it is a Step Two inquiry under *Free Enterprise*, the question is still one that can be resolved from the statutory text. The court is not importing later meanings into the text. Nor is it absolutely necessary for a court to consider extraneous information, such as legislative history, despite the fact that judges commonly elect to consider such material in the past when deciding whether to sever a provision.[168] Although it is occasionally more challenging to discern legislative purpose from statutory text, there are often statements of intent for various parts or for the statute as a whole.[169]

Teleological methodologies are those focused on the results achieved as a result of their utilization. The first is legal realism, an approach that arose in the 1920s and 1930s wherein judges step back from the words of the text, to "look[] beyond ideals and appearances" so as to ascertain what is "really going on" as a way of interpreting a statute's meaning.[170] The second is legal process, an approach developed in the 1950s and 1960s, which posits that every statute has a purpose, and so judges should envision themselves in a legislative setting and isolate the purpose Congress was pursuing in the statute, then reason through the accompanying rationale of such a purpose and interpret the statute in such a way as to advance that purpose.[171] An alternative name for this latter approach (or debatably a variation thereof) is purposivism, where, in more recent years, legal process has evolved to where the intermediary considerations of the legislative process are *per se* deemphasized

168. *See, e.g.*, *Chadha*, 462 U.S. at 932 (stating that when Congress makes its intentions clear through its actions, legislative history need not be considered when dealing with severability).

169. *See, e.g.*, Patient Protection and Affordable Care Act, Pub. L. No. 111-148 § 1501(a)(2)(D), 124 Stat. 119, 243 (2010), *as amended by* Health Care and Education Reconciliation Act of 2010, Pub. L. No. 111-152, 124 Stat. 1029 (2010) (providing that the purpose of the statute is to achieve "near-universal" health coverage).

170. BRIAN BIX, JURISPRUDENCE: THEORY AND CONTEXT 190 (5th ed. 2009).

171. Diarmuid F. O'Scannlain, *The Role of the Federal Judge Under the Constitution: Some Perspectives from the Ninth Circuit*, 33 HARV. J.L. & PUB. POL'Y 963, 967 (2010).

in favor of a direct examination of the public purpose supposedly served by the enactment at issue.[172]

Any of these approaches readily lend themselves to severability inquiries. As with the neutral inquiries, these teleological approaches are irrelevant in Step One of *Free Enterprise*, since courts are looking to the interdependency and interoperability of the statute's various provisions according to their textual terms. Under *Free Enterprise* Step Two, courts have myriad sources to consult under these teleological methods, from floor statements, to committee reports, to material from individual legislative offices. This enables a court to inquire into the reality of what (in the court's judgment) Congress was attempting to accomplish, and to invalidate a statute that does not achieve its objective without the discarded provision.

Thus severability can be applied regardless of the methodology any given jurist prefers. This is encouraging, in that often judgments vary as a result of conflicting approaches to statutory interpretation. A coherent doctrine governing severability can allow for that, and once a court reaches the stage of formulating a remedy, judges may at least be able to agree on how much of a statute to strike from the books. Such a statement may seem like the naïve optimism of one unaccustomed to jurisprudential reality, but hope springs eternal, no matter how counterfactual.[173]

---

172. *See* John F. Manning, *Exchange: What Divides Textualists from Purposivists?*, 106 COLUM. L. REV. 70, 71–73 (2006). This approach is best summarized by arguing that when the text of a statute varies with what the court finds to be the statute's purpose, "(1) Congress must have expressed its true intentions imprecisely, and (2) a judicial faithful agent [of Congress] could properly adjust the enacted text to capture what Congress would have intended had it expressly confronted the apparent mismatch between text and purpose." *Id.* at 72. Purposivism often arises in First Amendment analyses in law review literature, *see, e.g.*, Richard A. Posner, *Pragmatism Versus Purposivism in First Amendment Analysis*, 54 STAN. L. REV. 737 (2002); Jed Rubenfeld, *The First Amendment's Purpose*, 53 STAN. L. REV. 767 (2001), though its principles are not uniquely suited to any specific constitutional right, and this concentration may be a product of a scholarly debate that could also focus on other constitutional (or statutory) provisions.

173. Like any lawyer or scholar whose duties require reading many decisions of the Supreme Court and lower courts, I have no illusions that judges with contrary interpretive methodologies will often split in their judgments, with adherents of neutral schools often voting together against the conclusion agreed to by adherents of teleological schools. But my point here is that to the extent that there is a division of votes on the judgment, it will not stem from these schools, but rather from the judicial philosophy of the jurists. The philosophy held by a particular judge very often leads to adopting a particular school of interpretive thought. So while in theory different interpretive methodologies ought to be able to lead to the same conclusion regarding severing a given provision, the reality is that judges will still likely adjudicate statutes

## III. SEVERABILITY DOCTRINE IS BASED ON THREE SEPARATION-OF-POWERS PRINCIPLES

The Supreme Court declared in the recent case *Ayotte v. Planned Parenthood of Northern New England* that there are three principles underlying modern severability doctrine.[174] In a unanimous opinion by Justice Sandra Day O'Connor, the Court held:

> Three interrelated principles inform our approach to remedies. First, we try not to nullify more of a legislature's work than is necessary . . . . Second, mindful that our constitutional mandate and institutional competence are limited, we restrain ourselves from rewriting [a] law to conform it to constitutional requirements even as we strive to salvage it. . . . Third, the touchstone for any decision about remedy is legislative intent, for a court cannot use its remedial powers to circumvent the intent of the legislature.[175]

The Court then reiterated and reaffirmed these three principles in *Free Enterprise*.[176] We will explore each of these principles in detail below.

*Ayotte* was extremely helpful. Until 2006, only the Supreme Judicial Court of Massachusetts had provided a detailed explanation of how it formulated rules for severing invalid provisions from statutes.[177] Aside from that examination by a state court in 1854—in a decision subsequently adopted in substantial part by the U.S. Supreme Court in 1881[178]—case law provided dim illumination for the various severability rules employed. With *Ayotte*, the High Court has at long last provided a coherent rationale whereby lower courts and legal scholars can begin formulating a predictable framework. The Court's unanimity in *Ayotte* reflects an agreement spanning the full

---

based on each judge's conception of the proper judicial role in assessing public policy decisions within our tripartite constitutional framework.

174. 546 U.S. 320, 329–30 (2006). The facts and details in *Ayotte* are discussed in Part II.B.1.c, *supra*.

175. *Id.* (alterations and citations omitted).

176. *See* Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3161–62 (2010).

177. Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 98–99 (1854). This Article quotes the relevant passage in its entirety and discussed in Part II.A, *supra*. *See supra* notes 50–56 and accompanying text.

178. *See* Allen v. City of Louisiana, 103 U.S. 80, 83–84 (1881) (quoting *Warren*, 68 Mass. (2 Gray) at 99) (stating that if unconstitutional provisions are so connected with the general scope of the law, it may be impossible to strike them and still give effect to the intent of the legislature). *Allen* is likewise discussed in Part II.A.

spectrum of legal thought, and fortunately for the long-term development of severability doctrine it was also mercifully correct in its reasoning.[179] Beyond informing judicial decisionmaking, such a framework can also inform legislative efforts to design statutes that more easily lend themselves to resolving severability questions when provisions are found invalid.

### A. *Principle 1: Preference for Practical Minimalism*

The first severability principle from *Ayotte* is that the Court will "try not to nullify more of a legislature's work than is necessary . . . ."[180] The two operative words in this statement are "try" and "necessary." The first connotes that a court understands that severability is not an exact science, in that the court is deconstructing a document that was enacted as an integrated whole, and most often the ideas encapsulated by the invalid words have some bearing or connection with the words found in other provisions, which the court is attempting to salvage. The second does more than recognize the necessity for courts to strike down provisions that violate the Constitution. It also allows that this list of necessarily-rejected provisions includes more than the stand-alone unconstitutional provision. The Court is stating a preference for not striking down more of the statute than necessary, but does not limit that necessity to a provision that in isolation transgresses the Constitution. In carrying out its duty to nullify the inherently invalid provision—as it must—the task for a court becomes determining which companion provisions should fall with it, while acknowledging that this invalidating orbit extends beyond that which is intrinsically necessary. Otherwise stated, it goes beyond a judge inquiring, "What must I strike down?" to also include, "What else should I strike down with it?"

179. It also might be a tribute to Chief Justice John Roberts. Abortion is one of the most hotly-contested issues in American political life, and has dominated Supreme Court confirmations for more than a quarter-century. Roberts had only been Chief Justice for two months when *Ayotte* was argued on November 30, 2005. Yet when the Court's decision came down in January 2006, immediately before the retiring Sandra Day O'Connor was replaced by Samuel Alito, the Court handed down a unanimous decision on abortion written by O'Connor. This is either a testament to Roberts's ability to form consensus, or to O'Connor wanting to leave on a strong and unified note, or both.

180. *Ayotte*, 546 U.S. at 329.

The Court immediately explained the impetus for this first principle, "for we know that '[a] ruling of unconstitutionality frustrates the intent of the elected representatives of the people.'"[181] As the Court has said in reference to other juridical doctrines, the third branch's power must be limited in a democratic republic.[182] When invalidating programs created by statute, a "court must consider that all judicial interference with a public program has the cost of diminishing the scope of democratic governance."[183] As with the other doctrines referenced in Part II.B.2, *supra*, a consistent theme in juridical doctrines is recognition that unelected, life-tenured jurists are fundamentally unaccountable to the people,[184] and as such their power should be limited as an antidemocratic, countermajoritarian institution in a governmental system otherwise premised on the consent of the governed expressed through the political process.[185]

This severability principle arises from the limited role of unelected, unaccountable judges assigned to the courts by the Constitution in our democratic republic. Laws are the codification of public policy and are created through a process

181. *Id.* (quoting Regan v. Time, Inc., 468 U.S. 641, 652 (1984) (plurality opinion)).

182. *See* Lujan v. Defenders of Wildlife, 504 U.S. 555, 560 (1992) (quoting THE FEDERALIST NO. 48 (J. Madison)) (describing the uncertain boundaries of the judiciary's power); Valley Forge Christian Coll. v. Ams. United for the Separation of Church and State, Inc., 454 U.S. 464, 471 (1982) ("The judicial power of the United States . . . is not an unconditioned authority to determine the constitutionality of legislative or executive acts.").

183. Ill. Bell Tel. Co. v. Worldcom Techs., Inc., 157 F.3d 500, 503 (7th Cir. 1998); *accord* ALEXANDER M. BICKEL, THE LEAST DANGEROUS BRANCH 22 (2d ed. 1986) (quoting JAMES BRADLEY THAYER, JOHN MARSHALL 103–04, 106–07 (1901)); *see also* Michael D. Schumsky, *Severability, Inseverability, and the Rule of Law*, 41 HARV. J. ON LEGIS. 227, 278 (2004) (stating that the increasing willingness of the lawmaking branches to leave constitutional inquiries to the courts raises a serious concern of undermining the democratic process).

184. This is in reference to the fact that Article III federal judges hold their offices during "good behavior," U.S. CONST. art. III, § 1, cl. 2, which is generally understood to mean for life, absent some sort of felony or other notorious act that could realistically subject them to impeachment and removal by the U.S. House and Senate, respectively, *id.* art. I, § 2, cl. 5; *id.* § 3, cl. 6. Such occasions are rare, in that in the 222 years that judges have sat on the federal bench under Article III of the Constitution, only fifteen federal judges have been impeached, and only eight of those impeached judges have been removed. *See* FED. JUDICIAL CTR., HISTORY OF THE FEDERAL JUDICIARY, IMPEACHMENT OF FEDERAL JUDGES, http://www.fjc.gov/history/home.nsf/page/judges_impeachments.html (last visited July 9, 2011).

185. DaimlerChrysler Corp. v. Cuno, 547 U.S. 332, 341 (2006); *see also* THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776) (explaining that government is only established through the consent of the governed); ABRAHAM LINCOLN, GETTYSBURG ADDRESS, para. 3 (U.S. Nov. 19, 1863) (noting that there will government "by the people, for the people").

whereby the will of the American people can be expressed in a more direct fashion than is possible—or optimal—from the judiciary. A corollary of this democratic ideal is that when legislation is faulty, it is often preferable for the people's elected leaders to cure these deficiencies through the transparency and accountability of the political process. Consequently, the judiciary's role is to be the branch of last resort.[186] Beyond meaning that courts exercise judicial review only when other options to vindicate the rule of law have been exhausted, it also means that courts need to prefer using a scalpel to a sledgehammer when a situation presents multiple remedial options.[187]

An example of this first severability principle comes from the Supreme Court's most recent restatement of severability in *Free Enterprise*.[188] Having found an Appointments Clause violation in the tenure protections of Members of the new oversight board created by the Sarbanes-Oxley Act, the Court had to choose between two remedies to cure the constitutional infirmity. The first was to strike down the statute's provision creating the double for-cause tenure protections of the Board Members. The other was to strike down the words of enough of the provisions granting various powers to the Board to downgrade its Members until they would no longer have sufficient authority to be officers of the United States under the Appointments Clause. Between striking down one provision versus striking down dozens of provisions, Chief Justice John Roberts rejected the latter approach, reasoning that the Court cannot "blue-pencil" numerous, disjointed provisions of Sarbanes-Oxley because "such editorial freedom—far more extensive than our holding today—belongs to the Legislature, not the Judiciary. Congress of course remains free to pursue any of these options going forward."[189]

186. Nat'l Treasury Emps. Union v United States, 101 F.3d 1423, 1431 (D.C. Cir. 1996).

187. Justice Cardozo expressed this sentiment well while serving on the New York Court of Appeals. *See* People *ex rel.* Alpha Portland Cement Co. v. Knapp, 129 N.E. 202, 208 (N.Y. 1920) (Cardozo, J.) ("Our right to destroy is bounded by the limits of necessity. Our duty is to save, unless in saving we pervert.").

188. *See* Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138 (2010) (noting that the key to severability is legislative intent); *see also supra* Part II.B.1.d. (discussing the facts of *Free Enterprise*), *infra* Part IV.A (exploring the modern test for severability inquiries set forth by the Court in *Free Enterprise*).

189. *Id.* at 3162.

Professor Laurence Tribe raises the possibility that courts cannot sever an unconstitutional statute, as it would violate the Constitution's Bicameralism and Presentment Clauses.[190] Under the Constitution, that which becomes federal law must have been passed by both houses in identical form, then either signed by the President or re-passed by a two-thirds vote in both chambers of Congress in the event of a presidential veto.[191] The formalism required by these clauses has been reinforced by the Court as recently as 1998.[192] But this argument proves too much, as it categorically forecloses any possibility of severability in any instance, a premise that the judiciary has obviously categorically rejected. Assuming that the Court is unwilling to reverse course and completely jettison severability doctrine, Tribe's argument cannot gain much traction. It can serve as support for holdings of nonseverability, however, by providing a supplemental rationale encouraging courts to invalidate legislation in its entirety, as opposed to altering the statute so as to pass constitutional muster.

Another author objects that Tribe's argument fails because the statute in question was duly enacted pursuant to the Bicameralism and Presentment Clauses,[193] but that argument misses the point. Conceptually, a statutory enactment is a codification of a legislative bargain reduced to writing. Thus, this bargain is what Congress conceptually agreed to; this bargain is analogous to the "meeting of the minds" in contract law. Tribe thus makes a valid point. It just happens to be one that the Supreme Court has rightly concluded is inconsistent with the properly circumspect role of the judiciary in our tripartite form of government.

### B. *Principle 2: Courts Cannot Rewrite Statutes*

The Court in *Ayotte* explained its second principle of severability thus:

190. Laurence H. Tribe, *The Legislative Veto Decision: A Law by Any Other Name?*, 21 HARV. J. ON LEGIS. 1, 21–23 (1984); *See* New York *ex rel.* Hatch v. Reardon, 204 U.S. 152, 160 (1907). But it is critical to note that Holmes wrote during the era before the Supreme Court had articulated its first severability test, and before the Court set forth the presumptions that are central to the modern doctrine. *See supra* Part II.A & B. Recent decisions involving bicameralism and presentment include no language casting doubt on modern severability. *See, e.g.*, Clinton v. City of New York, 524 U.S. 417, 438 (1998).

191. U.S. CONST. art. I, § 7, cl. 2.

192. *See Clinton*, 524 U.S. at 427 n.12 (invalidating the Line Item Veto Act).

193. Nagle, *supra* note 5, at 228–29 & nn.129–30.

> Second, mindful that our constitutional mandate and institutional competence are limited, we restrain ourselves from "rewrit[ing] [a] law to conform it to constitutional requirements" even as we strive to salvage it. Our ability to devise a judicial remedy that does not entail quintessentially legislative work often depends on how clearly we have already articulated the background constitutional rules at issue. . . . But making distinctions in a murky constitutional context, or where line-drawing is inherently complex, may call for a "far more serious invasion of the legislative domain" than we ought to undertake.[194]

Simply restated, the second principle is that courts must refrain from effectively rewriting legislation under the rubric of judicial modesty. If an invalid provision cannot be severed without *de facto* rewriting a statute, then the statute should be entirely invalidated. Likewise a court cannot infer additional limiting language in a statute to confine it to constitutional boundaries, as a court cannot "dissect an unconstitutional measure and reframe a valid one out of it by inserting limitations it does not contain. This is legislative work beyond the power and function of the court."[195]

This principle includes several relevant components. The first is a reasoned admission of the limits the Constitution imposes on the judiciary. The limit is that a court's constitutional mandate does not extend to lawmaking. This limit is eminently reasonable, given the Court's accompanying admission that as an institution it has limited ability to write statutes competently. (Congress provides cynics with plenty of fodder to offer a snide remark that Congress evidently likewise has limited competence

194. *Ayotte*, 546 U.S. at 329–30 (quoting first Virginia v. Am. Booksellers Ass'n, Inc., 484 U.S. 383, 397 (1988), then United States v. Treasury Emps., 513 U.S. 454, 479 n.26 (1995)). Federal courts should apply this doctrine not only when examining federal enactments, but also when reviewing state statutes (illustrated by the fact that *Ayotte* was reviewing a New Hampshire statute). *See* Eubanks v. Wilkinson, 937 F.2d 1118, 1122–24 (6th Cir. 1991); Hill v. City of Hous., 789 F.2d 1103, 1112 (5th Cir. 1986), *aff'd*, 482 U.S. 451 (1987); Consumer Party v. Davis, 778 F.2d 140, 147 (3d Cir. 1985); *see also* Hynes v. Mayor of Oradell, 425 U.S. 610, 622 (1976). This rule may not apply, however, if there is a state court interpretation of the statute that narrows its scope to constitutional limits. *See* Smith v. Goguen, 415 U.S. 566, 575 (1974). This rule also extends to substatutory state authorities. *See Hill*, 789 F.2d at 1112, *aff'd*, 482 U.S. 451 (1987) ("The principles of federalism forbid a federal appellate court to arrogate the power to rewrite a municipal ordinance.").

195. Hill v. Wallace, 259 U.S. 44, 70 (1922). This is inconsistent with the doctrine of constitutional avoidance. *See supra* note 160 and accompanying text. Evidently just as severability trumps overbreadth doctrine, *see supra* note 158 and accompanying text, so too severability doctrine trumps constitutional avoidance.

when it comes to formulating statutes, but at least it is Congress's job to write laws, regardless of those statutes' effectiveness.) In explicating this principle, the Court expressly cited to one case, *American Booksellers*,[196] involving facial and overbreadth challenges to a law burdening the First Amendment, reinforcing as discussed in Part II.B.2 that severability doctrine shares a jurisprudential root with several other interpretive doctrines.

The second—and most important—component is that a court cannot engage in "quintessentially legislative work." The clear implication of this language is that it is possible for a court to usurp Congress's exclusive power under Article I to create legislation. To engage in such activity under the aegis of devising a remedy violates the separation of powers.[197] Since it cannot be denied that any judicial interpretation or construction imposed on a statute—or invalidating any one aspect of a statute—effectively reformulates a statute to some degree by imposing a new meaning upon a preexisting legal text, it becomes unavoidable that courts must draw the line at some point. This is, thus, a balancing requiring careful judicial reasoning and a degree of modesty.

This principle thereby explicitly states that it is possible for a court, under a facially-plausible claim of attempting to salvage a statute by severing an invalid provision, to essentially create a new statute. Such an activity is a "serious invasion of the legislative domain." The Framers understood the legislative power vested in Congress by Article I to be the means whereby legal duties would be imposed upon citizens, organizations, and governmental units, and legal rights declared and defined through the creation of statutes.[198] The judiciary would be

196. *Am. Booksellers*, 484 U.S. at 388–89.

197. This potential for interbranch conflict is unsurprising, given that whenever one branch of government is acting on the edge of the envelope on the purview of another branch, self-interest on the part of governmental actors in both implicated branches can create a result reminiscent of the reaction when one dog being walked down a street passes the yard of another dog. Growling and barking ensues, sometimes becoming sufficiently raucous as to become manifestly annoying to those unfortunate enough to be nearby at the moment. *Cf.* C. Vered Jona, *Cleaning Up for Congress: Why Courts Should Reject the Presumption of Severability in the Face of Intentionally Unconstitutional Legislation*, 76 GEO. WASH. L. REV. 698, 700 (2008) ("Throughout American history, the question of severability has reappeared repeatedly, reflecting an underlying tension between the courts and legislatures.").

198. 1 THE FOUNDERS' CONSTITUTION 382–87 (Philip B. Kurland & Ralph Lerner eds., 1987); *see also* INS v. Chadha, 462 U.S. 919, 952, 954 (1983); *see also, e.g.*, 1 RECORDS

invading Congress's domain, if severing invalid provisions from statutes would significantly alter any of these legal relationships.

The Supreme Court began developing this principle in severability as early as 1879, directing that:

> [W]hile it may be true that when one part of a statute is valid and constitutional, and another part is unconstitutional and void, the court may enforce the valid part where they are distinctly separable so that each can stand alone, it is not within the judicial province to give to the words used by Congress a narrower meaning than they are manifestly intended to bear in order that [the abridged statute is] within the constitutional power of that body . . . .[199]

This narrowing language sounds in the doctrine of constitutional avoidance, as discussed above, under which, when possible, a court declines to interpret statutes in a manner that would raise doubts regarding its constitutionality, thereby requiring a court unnecessarily to pronounce a constitutional holding.[200] By the aforementioned language, the *Ayotte* Court makes clear this narrowing track is not to be applied in cases where doing so would circumscribe severability, inserting a significant qualification into an otherwise generally-applicable doctrine of imposing narrowing constructions on statutes, which appears in a great many judicial inquiries.

Judicially rewriting a statute is the practical consequence of changing the import of a statute from a holistic perspective, with a court taking a gestalt of the statute's overall governing dynamics and with an eye to whether removing the invalid provision alters this quasi-artistic impression. "The presumption in favor of separability does not authorize the court to give the statute 'an effect altogether different from that sought by the measure viewed as a whole.'"[201] The reasoning for this principle

---

OF THE CONSTITUTIONAL CONVENTION (James Madison, June 21, 1787), *reprinted in* 2 THE FOUNDERS' CONSTITUTION, *supra*, at 29–32; 2 RECORDS, *supra*, at 25 (Madison, July 17, 1787), *id.* at 131, 151 (Committee of Detail), *id.* at 321 (Journal, Aug. 18, 1787), *reprinted in* 2 THE FOUNDERS' CONSTITUTION, *supra*, at 34–36; 1 JAMES KENT, COMMENTARIES 207–10 (1826), *reprinted in* 2 THE FOUNDERS' CONSTITUTION, *supra*, at 39–40. This characterization of the Framers' understanding of Congress's legislative role is also found in other accounts of the Constitution's adoption. *See, e.g.*, 1 RECORDS OF THE FEDERAL CONVENTION OF 1787, at 20, 45–46, 48, 54, 55, 57, 60, 61, 334, 349–50 (Max Farrand ed., rev. ed. 1966).

199. Trade-Mark Cases, 100 U.S. 82, 98 (1879).

200. *See supra* notes 160–63 and accompanying text.

201. Carter v. Carter Coal Co., 298 U.S. 238, 313 (1936) (quoting R.R. Ret. Bd. v. Alton R.R. Co., 295 U.S. 330, 362 (1935)).

also overlaps with the third and final principle, which is discussed below. "So here, to give the sections in question the effect suggested, it would be necessary to reject" Congress's intent for this statute.[202] "To do this would be to introduce a limitation where Congress intended none, and thereby to make a new . . . statute, which, of course, [courts] may not do."[203]

One law student articulated this principle well, writing:

> [S]triking down a bill instead of redrafting it protects the separation of powers contemplated by the Constitution and preserves a court's role as an adjudicatory rather than a legislative body.
>
> This is particularly important in cases where a court finds legislative knowledge of constitutional infirmities prior to enacting the law. If the legislature intended to pass a statute knowing it to be unconstitutional, a court would be overstepping its reactive role by proactively reshaping the statute, and possibly rewriting it, to cure the constitutional defect.[204]

The Court in *Free Enterprise* provided a classic illustration of this principle. Acknowledging that the Court could theoretically order different remedies to cure Sarbanes-Oxley's constitutional infirmity, Chief Justice Roberts wrote that the Court:

> [M]ight blue-pencil a sufficient number of the [Public Company Accounting Oversight Board]'s responsibilities so that its members would no longer be "Officers of the United States." Or we could restrict the Board's enforcement powers, so that it would be a purely recommendatory panel. Or the Board members could in future be made removable by the President, for good cause or at will. But such editorial freedom—far more extensive than our holding today—belongs to the Legislature, not the Judiciary.[205]

### C. *Principle 3: Remedies Should Be Consistent with Legislative Intent*

This proscription on rewriting statutes relates to the third principle in *Ayotte*, which the Court expounded as follows:

---

202. Butts v. Merchs.' & Miners' Transp. Co., 230 U.S. 126, 135 (1913).

203. *Id.*

204. Jona, *supra* note 197, at 712.

205. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3162 (2010).

> Third, the touchstone for any decision about remedy is legislative intent, for a court cannot "use its remedial powers to circumvent the intent of the legislature." After finding an application or portion of a statute unconstitutional, we must next ask: Would the legislature have preferred what is left of its statute to no statute at all? All the while, we are wary of legislatures who would rely on our intervention, for "[i]t would certainly be dangerous if the legislature could set a net large enough to catch all possible offenders, and leave it to the courts to step inside" to announce to whom the statute may be applied. "This would, to some extent, substitute the judicial for the legislative department of the government."[206]

In synthesizing this principle, the Court cited eight cases spanning 126 years wherefrom this proposition could be inferred.[207]

With its language that "it would certainly be dangerous if the legislature could set a net large enough to catch all possible offenders, and leave it to the courts"[208] to determine which transgressors can validly be sanctioned, the Court here cautioned Congress not to paint with strokes that are too broad when legislating. Congress cannot throw a legislative net covering vast areas—some of which are constitutional and others not—and essentially task the courts with determining which parts of the law can stand and which parts cannot. Congress must provide more definitive indications of its intent for a given statute—analogous perhaps in some way to the nondelegation doctrine[209]—to provide a baseline from which courts can begin judicial inquiries.

---

206. Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. at 330 (quoting first Califano v. Westcott, 443 U.S. 76, 94 (1979) (Powell, J., concurring in part and dissenting in part); third and second quotations United States v. Reese, 92 U.S. 214, 221 (1876)) (other citations omitted); *accord* INS v. Chadha, 462 U.S. 919, 931–32 (1983) (noting that the invalid portions of a statue are to be severed unless the Legislature would not have enacted those provisions which are within its power).

207. *Id.* (citing United States v. Booker, 543 U.S. 220, 227–29 (2005); Minnesota v. Mille Lacs, Band of Chippewa Indians, 526 U.S. 172, 191 (1999); Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 684 (1987); Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 234 (1932); Dorchy v. Kansas, 264 U.S. 286, 289–90 (1924); Employers' Liability Cases, 207 U.S. 463, 501 (1908); Allen v. City of Louisiana, 103 U.S. 80, 83–84 (1881); Trade-Mark Cases, 100 U.S. 82, 97–98 (1879)).

208. *Id.* (quoting United States v. Reese, 92 U.S. 214, 221 (1876)).

209. Congress cannot delegate legislative power to administrative agencies of the executive branch. Loving v. United States, 517 U.S. 748, 771 (1996). Under the nondelegation doctrine, Congress at minimum must therefore articulate some "intelligible principle" in a statute, and the actions of the implementing agency must be consistent therewith by only promulgating regulations consistent with that principle.

For the same reasons as the principle involving rewriting statutes, ascertaining intent is an inexact science. (Recall that the Court forthrightly noted these principles are interrelated.) Nonetheless, as Judge Richard Posner explains, "[I]nstitutions act purposively, therefore they have purposes. A document can manifest a single purpose even though those who drafted and approved it had a variety of private motives and expectations."[210] While individual floor statements or committee statements of rank-and-file members of Congress may be the least imputable to the whole body, statements of committee or party leaders may reflect more broadly-held views. The best such indicia of intent would be clear statements of purpose written into the text of the statute. Regardless of the specific sources cited as revelatory of purpose, this third principle requires some methodology for discovering this intent.

Context matters in legal interpretation as it does in so many other endeavors. "Interpretation of a word or phrase depends upon reading the whole statutory text, considering the purpose and context of the statute . . . ."[211] The Court has held that "statutory language cannot be construed in a vacuum. It is a fundamental canon of statutory construction that the words of a

---

Hampton & Co. v. United States, 276 U.S. 394, 409 (1928). According to the Court, this is because Congress's legislative role is fulfilled by articulating an intelligible principle in a statute, and agency regulations that are not pursuant to such a principle or derived from a statutory mandate thereby usurp Congress's constitutional role by becoming *de facto* statutes. *See* Field v. Clark, 143 U.S. 649, 192 (1892). The Supreme Court invoked this doctrine a couple times in the 1930s to invalidate parts of Franklin Delano Roosevelt's New Deal, *see* A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495, 537–38 (1935); Pan. Ref. Co. v. Ryan, 293 U.S. 388, 421, 430 (1935), but has not subsequently used it to strike down any statutes. The Court has since used the nondelegation doctrine as a rule of statutory construction, one that sounds in the doctrine of constitutional avoidance, *see supra* note 160. As an interpretative canon, a court will construe statutes if possible in such a way as to be pursuant to a statutory intelligible principle, if at all possible. *See* Whitman v. Am. Trucking Ass'ns, 531 U.S. 457, 474–76.

The practical consequence of the current state of the law is that agency regulations can go beyond procedural requirements implementing statutes to actually enact public policy, so long as those policies to do amount to "major policies." *See* Nat'l Petroleum Refiners Ass'n v. FTC, 482 F.2d 672, 675–83 (D.C. Cir. 1973), *cert. denied*, 415 U.S. 951 (1974). It should also be noted that some Justices regard the current state of the law as empowering executive agencies with more latitude than the Constitution allows. *See Whitman*, 531 U.S. 487–88 (Thomas, J., concurring).

210. Richard A. Posner, *Legal Formalism, Legal Realism, and the Interpretation of Statutes and the Constitution*, 37 CASE W. RES. L. REV. 179, 196 (1987). *But see* Frank H. Easterbrook, *Statutes' Domain*, 50 U. CHI. L. REV. 533, 547 (1983) ("It turns out to be difficult, sometimes impossible, to aggregate" the goals of individual legislators "into a coherent collective choice. Every system of voting has its flaws. The one used by legislatures is particularly dependent on the order in which decisions are made.").

211. Dolan v. U.S. Postal Serv., 546 U.S. 481, 486 (2006).

statute must be read in their context and with a view to their place in the overall statutory scheme."[212]

This resurfaces repeatedly in severability case law. For example, when the Court severed the legislative veto provision in *INS v. Chadha*, Members of Congress complained that they would never have delegated the authority in question to the President or his subordinate Attorney General without that power being checked by Congress's power to veto that authority.[213]

Specifically, the Supreme Court has instructed that interpreting terms in a statute "must, to the extent possible, ensure that the statutory scheme is coherent and consistent."[214] This does not mean that jurists must agree with the policies Congress is pursuing in the statute. "It is not [the Supreme] Court's task to decide whether the statutory scheme established by Congress is unusual or even bizarre."[215] When examining statutes—especially long and/or complex statutes—readers must get a gestalt of the statute and derive a sense of the overall scheme created by the statute. When a provision is found invalid and the court is devising a remedy, one that violates the legislature's intent will often be inconsistent with the legislative bargain embodied in the statute. This creates serious doubt as to whether the legislature would have passed such a statute.

The second and third principles from *Ayotte* overlap to such an extent that they sometimes seem to collapse into one inquiry. Some parts of judicial opinions seem to reflect both principles to such an extent that a reader could just as easily label it as arising from one as from the other. For example, which principle did the Supreme Court employ in setting forth the following method for making a severability inquiry? "If we should, in the case before us, undertake to make by judicial construction a law which Congress did not make, it is quite probable we should do what, if the matter were now before that body, it would be unwilling to do . . . ."[216] It implicates rewriting the statute as much as it does the statute's intent. Yet *Ayotte*

212. Davis v. Mich. Dep't of Treasury, 489 U.S. 803, 809 (1989).

213. Nagle, *supra* note 5, at 226.

214. Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 222 (2008).

215. Cuomo v. Clearing House Ass'n, L.L.C., 129 S.Ct. 2710, 2733 (2009) (Thomas, J., concurring in part and dissenting in part) (internal quotation marks and alterationss omitted).

216. Trade-Mark Cases, 100 U.S. 82, 99 (1879).

situates this "what-if" scenario in the context of the third severability principle, and so lawyers should argue such a hypothetical under the rubric of legislative intent. While this designation is questionable, it is where the Supreme Court has placed it, so the designation need not make sense. As others have said in various ways, the Supreme Court is not last because it is always right; we treat the Court as right because it is always last.[217]

Courts apply this third principle from *Ayotte* by engaging in an admittedly problematic creative exercise. As Justice George Sutherland wrote for the Court decades ago:

> The statutory aid to construction in no way alters the rule that in order to hold one part of a statute unconstitutional and uphold another part as separable, they must not be mutually dependent upon one another. Perhaps a fair approach to a solution of the problem is to suppose that while the bill was pending in Congress a motion to strike out the [invalid provision] had prevailed, and to inquire whether, in that event, the statute should be so construed as to justify the conclusion that Congress, notwithstanding, probably would not have passed the [remaining statute].[218]

While the mutual dependence in the first part of this passage is well within the realm of judicial discovery, some level of speculation is inextricable from the latter part of this sort of inquiry. While the exploration need not be fraught with peril in all circumstances—one can hypothesize scenarios under which a particular invalid provision appears central to the statutory scheme and thus a court would effectively rewrite the statute by retaining the statute without the provision, thereby contravening legislative intent—at minimum such hypothesizing requires cautious circumspection.

Yet such determinations are essential in severability inquiries, and so courts must venture where textualists fear to tread.[219] Many pieces of legislation are carefully-balanced packages, for

217. *See* Brown v. Allen, 344 U.S. 443, 540 (1953) (Jackson, J., concurring in judgment) ("We are not final because we are infallible, but we are infallible only because we are final.").

218. Carter v. Carter Coal Co., 298 U.S. 238, 313 (1936).

219. R.R. Ret. Bd. v. Alton R.R. Co., 295 U.S. 330, 362 (1935) (citing as a factor rejecting the severability of a statute that "it is 'unthinkable' and 'impossible' that the Congress would have created the" challenged program without the unconstitutional provisions).

which various provisions may be incidental or insignificant, but others may be vital. But for this third principle from *Ayotte* (and the second principle as well), courts could routinely excise central or even indispensible provisions from statutes, thereby rebalancing the legislation into something the legislature neither intended nor desired.

## IV. MODERN SEVERABILITY TEST IS A TWO-STEP ANALYSIS

The foregoing material provides the context and foundation for the modern test for severability. As explained in Part II.B, *supra*, the Supreme Court is now on its third version of the proper test for determining whether an invalid provision can be separated from the remainder of a statute to preserve the residuary. After 56 years of severance cases, the Court in 1932 developed its first clearly-adopted test in *Champlin*.[220] This test was then supplanted (but not overruled) by *Alaska Airlines* in 1987,[221] giving us the test that some authorities continue to erroneously regard as the current test. Instead, it is only part of the current test.

However, as explained in Part II.B.1.d, the Supreme Court recently modified severability doctrine in 2010 in the case *Free Enterprise Fund v. Public Co. Accounting Oversight Board*.[222] The Court thereby resegregated the *Alaska Airlines* test into what should henceforth be regarded as a two-prong test (as had previously the case with *Champlin*), wisely reversing the *Champlin* order to consider functionality before considering legislative intent.[223] Such a revised approach is especially advisable from

220. *See* Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 234 (1932); *see also supra* Part II.B.1. I use the term "clearly adopted" because there were test-like articulations from previous cases, *see, e.g.*, Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 564–65 (1902); Allen v. City of Louisiana, 103 U.S. 80, 83–84 (1881), subsequent courts would typically cite to one or more of these cases, but did not consistently cite to the same passage from the same case. By contrast, the cited page from *Champlin* became the commonly-cited provision in severability cases for more than half a century. *See, e.g.*, Buckely v. Valeo, 424 U.S. 1, 108–09 (1976).

221. *See* Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987); *see also supra* Part II.B.2.

222. 130 S. Ct. 3138 (2010).

223. *See id.* 3161–62 (discussing a two-step approach to severability). The U.S. Supreme Court should formally announce that severability is now a two-step test. However, it is also worth noting the Court hardly pioneered this approach in *Free Enterprise*. When a federal court is examining a state law under diversity jurisdiction, *see* 28 U.S.C. § 1332, severability is a question determined under state law, Burlington N. & Santa Fe Ry. Co. v. Doyle, 186 F.3d 790, 804 (7th Cir. 1999); *but see* Randall v. Sorrell, 548 U.S. 230, 262 (2006) (plurality opinion) (holding invalid provisions of a Vermont

the vantage point of those adhering to a textualist or originalist perspective,[224] as it reduces the significance of extratextual methodologies preferred by proponents of purposivism.[225] Although standing alone *Alaska Airlines* was by no means a failed test, making it central to the second step of the Court's revised two-step method is immensely superior.

One oft-cited case in severability analyses is *Hill v. Wallace,* wherein the Supreme Court invalidated a provision of the Future Trading Act pertaining to commodities trading as exceeding Congress's power under the Commerce Clause.[226] Writing for the Court, Chief Justice William Howard Taft expressly noted that there was a severability clause in the statute, then proceeded to say that the inherently invalid provision of the statute "is so interwoven with those regulations that they can not be separated. None of them can stand."[227] The Chief Justice went on to explain the rationale of such a conclusion, in saying of the severability clause, "Section 11 did not intend the court to dissect an unconstitutional measure and reframe a valid one out of it by inserting limitations it does not contain. This is legislative work beyond the power and function of the court."[228]

### A. *The Modern Test from* Free Enterprise

Without rehashing the discussion of the facts of *Free Enterprise* from Part II.B.1.d, recall that the *Free Enterprise* case was a challenge to the Sarbanes-Oxley Act, in which the Court held a tenure-protection provision of the statute unconstitutional, and then proceeded to determine whether the offending provision

---

campaign finance law nonseverable, but giving no indication that the Court was applying Vermont severability doctrine). The evolution of severability into a two-step inquiry in which functionality is examined before legislative intent is also seen in at least some state courts. For example, the Indiana Supreme Court held six years prior to *Free Enterprise* that:

> A statute bad in part is not necessarily void in its entirety. Provisions within the legislative power may stand if separable from the bad. But a provision, inherently unobjectionable, cannot be deemed separable unless it appears both that, standing alone, legal effect can be given to it and that the legislature intended the provision to stand, in case others included in the act and held bad should fall.

State v. Barker, 809 N.E.2d 312, 317 (Ind. 2004) (quoting Dorchy v. Kansas, 264 U.S. 286, 289–90 (1924)).

224. *See supra* notes 165–67 and accompanying text.

225. *See supra* notes 170–72 and accompanying text.

226. 259 U.S. 44, 66–70 (1922).

227. *Id.* at 70.

228. *Id.*

could be severed from the remainder of this lengthy statute. Writing for the Court, Chief Justice John Roberts restated the framework governing severability as follows:

> Generally speaking, when confronting a constitutional flaw in a statute, we try to limit the solution to the problem, severing any problematic portions while leaving the remainder intact. Because the unconstitutionality of a part of an Act does not necessarily defeat or affect the validity of its remaining provisions, the normal rule is that partial, rather than facial, invalidation is the required course . . . .
>
> The Sarbanes-Oxley Act remains fully operative as a law with these tenure restrictions excised. We therefore must sustain its remaining provisions unless it is evident that the Legislature would not have enacted those provisions . . . independently of that which is invalid.[229]

This should be understood as creating a sequential two-step framework for conducting severability inquiries, and severability analyses should follow what they explicitly label as a two-step approach.

### 1. *Free Enterprise* Step One: Functionality

The first step in a severability inquiry is determining whether the statute is still literally functional without the invalid provision.[230] Specifically, the statute must continue to be "*fully* operative as a law,"[231] as opposed to merely functioning partially or haltingly. Thus this prong goes beyond asking if the abridged statute is effectively dead, instead asking if severing the invalid provision renders the remaining statute lame.

The statute must be fully functional, not merely substantially functional or mostly functional. Elsewhere this Article shows aspects of severability doctrine relevant to contract law. To imperfectly analogize to the law of contracts, the functionality step requires the U.C.C. standard of strict compliance for the

229. *Free Enter.*, 130 S. Ct. at 3161 (internal quotation marks and citations omitted) (second ellipsis in the original).

230. *Id.*

231. *Id.* (quoting New York v. United States, 505 U.S. 144, 186 (1992)) (emphasis added).

sale of goods, not the common-law rule of substantial performance.[232]

This requirement has fortunately been expounded over time. The Supreme Court long ago instructed that a court is not permitted "to reject a part which is unconstitutional and retain the remainder, [if] it is not possible to separate that which is unconstitutional . . . from that which is not."[233] By itself, such a rule begs the question of how a judge determines if it is possible to separate the offending provision. A century of cases since that 1906 command have subsequently elaborated on this statement, enabling courts to specifically consider whether any other provisions of a statute are incapacitated by the excision of the invalid provision, as examined throughout this Article.

The *Free Enterprise* Court significantly improved severability doctrine by asserting functionality before intent.[234] It allows a court to use a four-corners approach of looking to the various remaining sections, asking for *each* provision whether it can *fully* function without the invalid provision. In exploring possible answers, a judge has the full range of rules and cannons of statutory construction that courts employ on a daily basis. This doctrinal modification answers the criticism of some scholars that severability doctrine requires leaning too heavily on debatable indicia of legislative intent,[235] enabling courts to employ a more text-centric approach consistent with modern judicial trends.[236] Consequently, it should never be necessary to consult any extraneous sources in finding whether Step One is

232. *See, e.g.*, Pomerantz Paper Corp. v. New Cmty. Corp., 25 A.3d 221, 231–32 (N.J. 2011); Astor v. Boulos Co., 451 A.2d 903, 906 (Me. 1982); T.W. Oil, Inc. v. Consol. Edison Co., 443 N.E.2d 932, 937 (N.Y. 1982) (discussing U.C.C. § 2-508).

233. Ill. Cent. R.R. Co. v. McKendree, 203 U.S. 514, 529 (1906).

234. Although this is the Court's first holding to that effect, the Court had suggested such an ordering of questions even before *Champlin*. *See* Dorchy v. Kansas, 264 U.S. 286, 290 (1924) (dictum) (stating that "a provision, inherently unobjectionable, cannot be deemed separable unless it appears *both* that, standing alone, legal effect can be given it and that the legislature intended the provision to stand, in case others included in the act and held bad should fall") (emphasis added).

Professor Cass Sunstein commented in 1989 that text-centric interpretive methods were experiencing a renaissance. *See* Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 410 n.16 (1989). Increasing focus on text in the 1980s notwithstanding, his statement was premature in 1989, and in 2011 it still remains to be seen whether these approaches are ascendant.

235. *See* Movsesian, *supra* note 27, at 42; Dorf, *Facial Challenges*, *supra* note 150, at 291; Nagle, *supra* note 5, at 206 (discussing the difficulties of determining legislative intent).

236. *See generally* William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. REV. 621, 623–24, 656–67 (1990) (discussing Justice Scalia's originalist approach of statute interpretation and its influence on the court).

satisfied, as the court only considers whether Provision A is somehow dependent by reference or inference on an invalid Provision B, such that some aspect of Provision A is linguistically or logically incapacitated or rendered nonsensical—or functionally incapacitated—without Provision B.

One major severability case from decades past illustrates how Step One should be applied. A 1922 case marked one of the rare occasions that the Supreme Court struck down an entire statute on account of a single invalid provision despite the statute containing a severability clause.[237] In *Hill v. Wallace*, the Court invalidated the entirety of the Future Trading Act, holding that notwithstanding its severability clause,[238] the statute's other operative provisions were so interwoven with the invalid provision that they could not functionally be separated.[239]

At least one prominent scholar suggests that a lack of functionality should be the only grounds for finding an unconstitutional provision nonseverable.[240] But the Supreme Court rejects this argument, as well it should. The judicial role under such circumstances is not merely a mechanistic analysis of grammar and syntax. Rather, the Supreme Court's explication of the judicial power also assigns the judiciary the role of considering whether Congress's will is being frustrated when the court removes an invalid provision.

### 2. *Free Enterprise* Step Two: Legislative Intent and the Legislative Bargain

If the individual provisions of the statute are still operable without the invalid provision, then the court turns to the question of legislative intent. The court "must sustain its remaining provisions '[u]nless it is evident that the Legislature would not have enacted those provisions . . . independently of that which is [invalid].'"[241] Courts are to imagine that Congress was faced with a bill containing the statute minus the invalid provision, and determine whether Congress would still have

237. Jona, *supra* note 197, at 702.

238. *See* 259 U.S. 44, 70 (1922).

239. *Id.* at 70–72.

240. *See* 1 LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW 142 n.11 (3d ed. 2000).

241. Free Enter. Fund v. Pub Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3151, 3161 (2010) (quoting New York v. United States, 505 U.S. 144, 186 (1992)) (alterations in the original).

voted in favor of the bill.[242] The Supreme Court explained in 1909:

> "It remains to inquire whether it is plain that Congress would have enacted the legislation had the act been limited to" [its effect without the invalid provision] . . . . If we are satisfied that it would not, or that the matter is in such doubt that we are unable to say what Congress would have done omitting the unconstitutional feature, then the statute must fall.[243]

In 1936 the Court added to this reasoning, holding that several presumably-valid statutory provisions are "so related to and dependent upon the [invalid] provisions . . . as to make it clearly probable that the latter being held bad, the former would not have been passed. The fall of the latter, therefore, carries down with it the former."[244]

### a. *Ascertaining Legislative Intent*

The question cannot be whether the statute sans the invalid provision fully effectuates legislative intent, since by definition the invalidation of any provision which Congress adopted thwarts some aspect of legislative intent.[245] In answering this hypothetical inquiry, the test is "whether the statute will function in a *manner* consistent with the intent of Congress."[246] This requires a court to find Congress's intent in enacting the statute, which cannot be an exact science but which nonetheless must be attempted and then applied to the case. Determining intent requires interpreting any provision in the context of the statute as a whole,[247] which means that the principles invoked in Step Two overlap to some degree with those relevant to Step One.

Subsequent severability cases showcase this caveat. The *New York* Court opined that "where Congress has enacted a statutory

242. Carter v. Carter Coal Co., 298 U.S. 238, 312–13 (1936). A severability clause "provides that in the event that the original law is held partly invalid, a fallback of the original law minus the invalid provision or application will take effect." Dorf, *Fallback Law*, *supra* note 29, at 305.

243. El Paso & Ne. Ry. Co. v. Gutierrez, 215 U.S. 87, 97 (1909).

244. *Carter*, 298 U.S. at 316 (citing Int'l Textbook Co. v. Pigg, 217 U.S. 91, 113 (1910); Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 98–99 (1854)).

245. Hayes, *supra* note 83, at 141.

246. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987) (emphasis in original).

247. Dolan v. U.S. Postal Serv., 546 U.S. 481, 486 (2006) (holding that interpreting statutory language "depends upon reading the whole statutory text, considering the purpose and context of the statute, and consulting any precedents or authorities that inform the analysis").

scheme for an obvious purpose, and where Congress has included a series of provisions operating as incentives to achieve that purpose, the invalidation of one of the incentives should not ordinarily cause Congress's overall intent to be frustrated."[248]

Judges are not to make this requirement insuperable. Jurists must not say, "We do not see overwhelming evidence that Congress would have voted for this bill without the invalid provision, so we will sever it."[249] Such a standard would almost require challengers to disprove a negative, since unless Congress includes an express *nonseverability* provision, or concurrently passes a joint resolution proclaiming that it would not have enacted the preceding legislation without every provision in the bill (even suggesting such a thing would be fanciful), it would be almost impossible to present irrefutable proof that the invalid provision was intended to be fatal. Instead, courts must instead determine whether, "eliminating invalid parts, the Legislature would have been *satisfied* with what remained."[250]

A court must "seek to determine what 'Congress would have intended' in light of the Court's constitutional holding."[251] Heading off any ambiguity in this holding, the *Booker* Court quoted a plurality opinion from 1996, asking the question: Would Congress still have passed the remainder of the statute if Congress had known the invalid provision would be struck down?[252] This reconciles the *Booker* holding with existing severability case law, making it evident that the Court was not modifying the rule for severability.

The Supreme Court in *Booker* held that in choosing between options, a court must select the severability remedy that is "more compatible with the Legislature's intent as embodied" in the statute,[253] the one that "would deviate less radically from

248. New York v. United States, 505 U.S. 144, 186 (1992).

249. Ironically (and surprisingly) as this Article was being edited, the Eleventh Circuit severed a provision of controversial legislation with an opinion that applied this unworkable and incorrect standard. *See* Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health and Human Servs., 648 F.3d 1235, 1322–23 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

250. Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 235 (1932) (emphasis added).

251. United States v. Booker, 543 U.S. 220, 246 (2005) (quoting Denver Area Educ. Telecomms. Consortium, Inc. v. FCC, 518 U.S. 727, 767 (1996) (plurality opinion)).

252. *Id.* (quoting *Denver Area Telecomms. Consortium*, 518 U.S. at 767).

253. *Booker*, 543 U.S. at 246.

Congress's intended system."[254] Justice Stephen Breyer wrote this part of the opinion for the Court,[255] and limited the Court's holding with the caveat, "In [this] context—a highly complex statute, interrelated provisions, and a constitutional requirement that creates fundamental change—we cannot assume that Congress, if faced with the statute's invalidity in key applications, would have preferred to apply the statute in as many other instances as possible."[256] Although decided five years prior to *Free Enterprise*, Breyer made clear that this reasoning applied in what this article labels *Free Enterprise* Step Two, adding, "Neither can we determine likely congressional intent mechanically. We cannot simply approach the problem grammatically, say, by looking to see whether the constitutional requirement and the words of the Act are linguistically compatible."[257] (Such an approach would instead fall within *Free Enterprise* Step One.) Breyer also adds the caveat that remaining independent provisions are to be severed if they are also "consistent with Congress's basic objectives in enacting the statute."[258]

*Free Enterprise* also contains additional language suggesting courts are not to set a dauntingly-high bar when examining legislative intent. Chief Justice Roberts specifically reasoned, "Because the unconstitutionality of a part of an Act does not necessarily defeat *or affect* the validity of its remaining provisions, the normal rule is that partial . . . invalidation is the required course . . . ."[259] Two points relevant to this discussion must be gleaned from this passage. First, the Court is broadening the orbit of severability beyond those provisions that would be substantially *defeated* by removing the unconstitutional provision to also include those that would be *affected.*

254. *Id.* at 247.

255. *Id.* at 248. Breyer wrote the second part of the majority opinion, and was joined by Chief Justice Rehnquist and Justices O'Connor, Kennedy, and Ginsburg. The first part of the majority opinion was written by Justice John Paul Stevens, joined by Justices Scalia, Souter, Thomas, and (again) Ginsburg. *See id.* at 226.

256. *Id.* at 248.

257. *Id.*

258. *Id.* at 259 (citing Regan v. Time, Inc., 468 U.S. 641, 653 (1984) (plurality opinion)). Although *Free Enterprise* formally recast the severability test by building upon *Alaska Airlines*, *Booker* presaged the resegregation of severability analyses, enumerating three elements that must be met for an invalid provision to be severable. A court "must retain those portions of [an] Act that are (1) constitutionally valid, (2) capable of functioning independently, and (3) consistent with Congress' basic objectives in enabling the statute." *Id.* at 258–59 (internal quotation marks and citations omitted).

259. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3161 (2010) (emphasis added) (internal quotation marks and citations omitted).

This latter language is what is necessary to reconcile *Free Enterprise* with cases from throughout the previous 156 years. The language in the first severability case from a state court, which as we have seen was later substantially adopted by the Supreme Court, is that between a valid and an invalid provision, they cannot be "connected with and dependent on each other."[260] The provisions are severable "*if* the parts are *wholly* independent of each other,"[261] adding that "all the provisions which are thus dependent, conditional, or connected" must fall with the invalid provision.[262] This trio of words indicates an intent to include any form of significant relationship between statutory provisions. It is what the Court meant when it held that if an invalid provision "is of such import that the other sections without it would cause results not contemplated or desired by the legislature, then the entire statute must be held inoperative."[263] "To do this would be to introduce a limitation where Congress intended none, and thereby to make a new . . . statute, which, of course, [courts] may not do."[264] After all, even a severability clause "in no way alters the rule that in order to hold one part of a statute unconstitutional and uphold another part as separable, they must not be mutually dependent upon one another."[265] Otherwise stated, courts are to use this standard to determine "whether the statute will function in a *manner* consistent with the intent of Congress."[266] To do so would be to "rewrite [a] law to conform it to constitutional requirements" and thereby "to circumvent the intent of the legislature."[267]

Roberts's use of the word "validity" in the phrase "defeat or affect the validity of its remaining provisions" must be defined, because it does not refer to the inherent unconstitutionality of those remaining provisions. It was not the best choices of words (which is rare for a Chief Justice who has consistently demonstrated clear and unambiguous language in his opinions), and must be read in context. In a severability inquiry,

260. Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 99 (1854).

261. Allen v. City of Louisiana, 103 U.S. 80, 84 (1881) (emphasis added).

262. *Id.* at 84 (quoting *Warren*, 68 Mass. (2 Gray) at 99).

263. Connolly v. Union Sewer Pipe Co., 184 U.S. 540, 565 (1902).

264. Butts v. Merchs.' & Miners' Transp. Co., 230 U.S. 126, 135 (1913).

265. Carter v. Carter Coal Co., 298 U.S. 238, 313 (1936).

266. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987) (emphasis in original).

267. Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 329–30 (2006) (alterations and citations omitted).

once a constitutionally invalid provision is identified, the question becomes whether that provision can be separated from constitutionally valid provisions. So in the sense of constitutionality, all the remaining provisions are "valid" and this intrinsic validity cannot be altered by the invalid provision.

Instead, "valid" in this passage should be understood as referring to two things. The first would be the continued functionality of those statutes under Step One. The second arises if a court is in *Free Enterprise* Step Two, and would be whether those provisions still impact the overall statutory scheme in a manner consistent with Congress's intent—that its "validity" in this sense is its continuing to occupy the space Congress assigned to it and fulfill the role that Congress envisioned for it as a component of the legislative bargain. Such a reading is necessary to harmonize this *Free Enterprise* passage with the core holding of *Alaska Airlines*,[268] a holding which the *Free Enterprise* Court gave no indication it wished to disturb.

At least one scholar—Professor Larry Tribe—is critical of an intent-centric approach to severability,[269] which is central to a court's inquiry when in Step Two. And a plausible argument for judicial modesty and self-restraint can be made that courts should confine severability inquiries to accessing the ongoing functionality of the shortened statute without musing about collective legislative intent. But as discussed throughout this Article, courts and other scholars have instructed how legislative intent should be discerned, and equally-plausible arguments can be made that there is no modesty in insisting on applying only part of an enactment when the modified legislation no longer serves the goals driving its adoption, and may in fact now work against those goals.

The Supreme Court's most recent application of this inquiry is at least somewhat helpful. When proceeding to Step Two in *Free Enterprise*, Chief Justice Roberts first noted that the answer to the question of whether Congress would have passed the

268. *See Alaska Airlines*, 480 U.S. at 685 ("The more relevant inquiry in evaluating severability is whether the statute will function in a manner consistent with the intent of Congress.") (emphasis omitted).

269. *See* LAURENCE H. TRIBE, CONSTITUTIONAL CHOICES 79–83 (1985) (noting that intent-based approaches raise questions as to how courts can enforce laws that were never duly enacted).

abridged statute can be "elusive,"[270] quoting *INS v. Chadha*, though the Court in *Chadha* expanded upon that point to clarify that the inquiry is elusive only in cases where there is no severability clause in the statute,[271] (as was the case with Sarbanes-Oxley),[272] and also suggested this proposition four years later in *Alaska Airlines*.[273] This elusiveness should thus only attend statues lacking a severability clause. The final part of the *Free Enterprise* Court's reasoning illustrates this approach, as the Court then observed that "nothing in the statute's text or historical context makes it 'evident' that Congress" would rather have no statute than the truncated statute.[274] On this reasoning, the Court concluded that the intent prong was satisfied.

The Court's reasoning in those passages amply demonstrates the value of a severability clause for ascertaining legislative intent. In *Chadha*, the Court determined that "we need not embark on that elusive inquiry since Congress itself has provided the answer to the question of severability in § 406 of the Immigration and Nationality Act" by including a severability clause.[275] Legislative wrangling between various politicians, each with a different agenda, creates a morass that must be waded through for those intrepid lawyers seeking to divine an intelligible institutional intent from Congress's actions. Chief Justice Warren Burger was being quite charitable to Congress when he said that finding Congress's intent is "elusive" without such a clause, a charity perhaps augmented by his adherence to the ideal that when a Chief Justice writes for the Court, he should be particularly gracious when speaking critically of a

270. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3161 (2010) (quoting INS v. Chadha, 462 U.S. 919, 932 (1983)).

271. *Chadha*, 462 U.S. at 932.

272. Hans Bader, Free Enterprise Fund v. PCAOB*: Narrow Separation-of-Powers Ruling Illustrates That the Supreme Court Is Not "Pro-Business"*, 2009–2010 CATO SUP. CT. REV. 269, 279 (2010); *see also* Plaintiffs' Reply Memorandum Supporting Motion for Summary Judgment at 27–28 (Dkt. 47), *Free Enter.*, 2007 U.S. Dist. LEXIS 24310 (D.D.C. Mar. 21, 2007) (No. 1:06-cv-217-RMU); *cf.* Brief for Respondent Pub. Co. Accounting Oversight Bd. at 48–49, *Free Enter.*, 130 S. Ct. 3138 (No. 08-861).

273. *See Alaska Airlines*, 480 U.S. at 686.

274. *Free Enter.*, 130 S. Ct. at 3162 (quoting *Alaska Airlines*, 480 U.S. at 684).

275. *Chadha*, 462 U.S. at 932. That provision reads, "If any particular provision of this Act, or the application thereof to any person or circumstance, is held invalid, the remainder of the Act and the application of such provision to other persons or circumstances shall not be affected thereby." *Id.* (quoting note following 8 U.S.C. § 1101).

coequal branch of the federal government.[276] Severability clauses are so helpful in regard to assessing Congress's intent that it would be beneficial if Congress would make it a requisite part of legislation to include either an express severability clause or (here is a thought to take to heart, not that Congress would do so frequently) an express nonseverability clause in much of its legislation.[277]

This intent aspect of a severability inquiry has important implications for formulating national policy in legislation. Consider the result in the context of legislative theory. Legislation is proposed, codified in a bill. The bill is introduced by its sponsor. It is referred to a committee. There are hearings with testimony, following which the committee issues reports. After hearings the committee has a mark-up session, where amendments are proposed. The bill is debated, then voted out of committee. It goes to the floor of the full chamber, where there is additional debate and further opportunities to offer amendments. It then passes the chamber. This entire process is then repeated in the counterpart legislative chamber. For those few bills that make it through the entire process, the President either signs the bill into law, or vetoes it.

Every stage of this process involves negotiation and discussions between the political actors. Each of these is an expression of

276. One shudders to think what sort of adjective Justice Antonin Scalia would substitute for "elusive," and perhaps we will have occasion to find out, likely to the amusement of legal media commentators and the consternation of Members of Congress. Justice Scalia has a habit of using sarcasm and acerbic verbiage in his opinions. For example, in a 2006 habeas corpus case, a fractured Court invalidated the President's system for processing terrorist detainee suspects without statutory authority. In his opinion, Justice Stephen Breyer says: "Nothing prevents the President from returning to Congress to seek the authority he believes necessary." Hamdan v. Rumsfeld, 548 U.S. 557, 636 (Breyer, J., concurring). When Congress subsequently passed such legislation and then the Court struck it down, Justice Scalia said of his colleagues: "Turns out they were just kidding." Boumediene v. Bush, 553 U.S. 723, 831 (2008) (Scalia, J., dissenting). Other times he can respond to arguments he rejects with a single provocative word. *See, e.g.*, District of Columbia v. Heller, 554 U.S. 570, 587 (2008) ("Grotesque.").

Such statements routinely amuse his admirers and irritate his critics. Those of us in the legal community whose duties require us to frequently attend Supreme Court arguments can attest that Scalia is equally entertaining during oral argument, with exchanges closely resembling the way a tough law professor conducts a classroom discussion.

277. Recent literature includes several discussions of nonseverability clauses. *E.g.*, Noah, *supra* note 26, at 238–39; *see also* ABNER J. MIKVA & ERIC LANE, AN INTRODUCTION TO STATUTORY INTERPRETATION AND THE LEGISLATIVE PROCESS 167–68 (1997) (stating that nonseverability clauses which tie certain provisions of legislation to one another may help protect against the undermining of legislative compromises); Israel E. Friedman, Comment, *Inseverability Clauses in Statutes*, 64 U. CHI. L. REV. 903, 903 & n.4, 907 (1997).

individual intent, which is aggregated into what we conceptualize as "legislative intent." Negotiated bargains are ubiquitous throughout this entire narrative, and it is a common occurrence that any particular provision that results from one of these bargains is a *sine qua non* but for which the legislation would not have progressed toward final enactment.

Without the intent prong in severability analysis—the second step of *Free Enterprise*—legislation could be retained that no longer worked the people's will as expressed by their elected political representatives through the democratic process. Legislation is, after all, formulated to deliberately serve a particular public purpose, and so removing part of a statute should cast the statute's continuance in doubt unless a court is satisfied that its overarching purpose is still advanced. This explains why various courts have stated that severability is ultimately a question of legislative intent.[278] As one treatise states, when the "purpose of [a] statute is defeated by the invalidity of part of the act, the entire act is void."[279] This hearkens back to the Supreme Court's first severability test from 1881, asking "whether the unconstitutional provisions are so connected with the general scope of the law as to make it impossible, if they are stricken out, to give effect to what appears to have been the intent of the legislature."[280]

While indicia of legislative intent is less precise than rules of statutory interpretation, and practices that consult extraneous material do not enjoy the same universal acceptance as principles of textual construction, plumbing the depths of legislative intent need not always be a fool's errand or one that will inexorably lead a judge to the conclusion he was already predisposed to reach. There are various principles for finding such intent. Moreover, in ascertaining legislative intent, courts can employ the various rules utilized in contract law to find the intent of the parties.[281] This is all the more appropriate because severance in statutory remedies is similar to the severability rule in contracts, where a court can sever a term if the parties would

278. *See, e.g.*, Ala. Power Co. v. United States Dep't of Energy, 307 F.3d 1300, 1307 (11th Cir. 2002); New Haven v. United States, 809 F.2d 900, 907 (D.C. Cir. 1987).

279. NORMAN J. SINGER, SUTHERLAND STATUTORY CONSTRUCTION § 44.07 (4th ed. 1986).

280. *Id.*

281. Movsesian, *supra* note 27, at 43.

still have formed the contract without the invalid provision,[282] as explained in the *Restatement.*[283]

b. *Adhering to the Legislative Bargain*

Courts have been authoritatively declaring congressional intent and then formulating holdings pursuant to those declarations throughout severability case law. When examining a claim under the Civil Rights Act of 1875, the Supreme Court in 1913 began its severability analysis by noting Congress's "manifest purpose was to enact a law which would have a uniform operation wherever the jurisdiction of the United States extended."[284] When the Court held an invalid provision severable in 1984, Justice Byron White wrote for a plurality of the Court in *Regan v. Time* that the "policies Congress sought to advance" would still be effectuated by the remaining statute, thus achieving "the main purposes" of the statute.[285]

The many examples found in case law provide a considerable amount of guidance on how courts are to find whether Congress's intent for a statute can be preserved without the invalid provision, and if so a court can sever that provision and retain the residuary. Statutes are bundles of legislative concerns and priorities. Some have primacy, but for which the quintessence of the statute would be modified. Others are of secondary or even tertiary significance, the diminution of which does not subvert Congress's overarching design for the statute.

The Court in *New York* signaled that severability is more likely when the invalid provision is merely an aid to the "main purpose" of a statute,[286] suggesting that a court should refrain from severing in situations where excising the offensive

282. *Id.* at 43–44. In this respect, to analogize a severable provision from a nonseverable one, a nonseverable statutory provision is likely a condition (a necessary provision) in a contract, while a severable provision is just a term (an unnecessary provision). Conditions are provisions the violation of which constitutes a breach of the contract—a "deal breaker"—which cancels the obligations of the other party and so the contract is thereby undone. So too, the invalidity of a nonseverable provision nullifies the entire statute.

283. *See* RESTATEMENT (SECOND) OF CONTRACTS § 184(1) (1981).

284. Butts v. Merchs.' & Miners' Transp. Co., 230 U.S. 126, 133 (1913).

285. 468 U.S. 641, 653–55 (1984) (plurality opinion).

286. New York v. United States, 505 U.S. 144, 186–87 (1992) (quoting Reagan v. Farmers' Loan & Trust Co., 154 U.S. 362, 396 (1894) (holding of penalties that are "simply in aid of the main purpose of the statute" that "[t]hey may fail, and still the great body of the statute have operative force, and the force contemplated by the legislature in its enactment")).

provision would undermine the statute's manifest purpose. Applying *Alaska Airlines* as an example of severing an unconstitutional provision which was unshielded by a severability clause, the *New York* Court held that the invalid provision could be severed both because the remainder of the statute "is still operative *and* it still serves Congress'[s] objective . . . ."[287] Thus post-*Alaska Airlines*, courts place revived the judiciary's early emphasis on considering the purpose of the statutory scheme in assessing whether a given provision is separable. The Court reinforced this aspect of its holding by concluding, "The purpose of the Act is not defeated by the invalidation of the [unconstitutional provision], so we may leave the remainder of the Act in force."[288] Conversely, the closer the proximity of the invalid provision to the statute's center of gravity, the more likely a court should not preserve the remaining statute. Various lower federal courts have likewise understood severability doctrine in this fashion, such as assessing whether the invalid provision is "central" to the "core mission" of the statute.[289]

The presumption of severability can also be included in this framework, at least when a challenged statute contains a severability clause. If severability is presumed, an invalid provision is presumed severable because it is presumed to be of minor significance to the legislative bargain. The burden often rests on the challenger to persuade the court that this particular unconstitutional provision is instead of major significance. Such provisions are critical to Congress's legislative bargain, going to the heart of the statutory scheme created by the enactment. Few provisions can be shown to be so integral to Congress's plan. This will keep instances of total invalidation rare, preserving the norm of partial invalidation.

287. *Id.* at 187 (emphasis added).

288. *Id.*

289. *See, e.g.*, McConnell v. FEC, 251 F. Supp. 2d 176, 435 (D.D.C. 2003) ("The provisions I have found unconstitutional are all provisions . . . that are not central to [the statute's] core mission and are entirely severable without doing injustice to the remainder of the law."); *id.* at 776 (asserting that a severability clause "does not relieve this Court of its obligation to determine if the [remaining provisions] can stand alone, and if Congress would have enacted [the statute] knowing [these provisions] would be held unconstitutional").

### c. *Provisions of Major Significance are Nonseverable*

In sum, under *Free Enterprise* Step Two courts must properly conceptualize Congress's legislative bargain embodied in the enacted legislation, and then hold nonseverable any invalid provision significant to that deal—reaching to its core, because the statute cannot function in the manner Congress intended without that provision.

Nonseverable provisions are those that are so significant that they upset the legislative bargain discussed above. They need not render other provisions incoherent or nonsensical, as such matters fall within Step One. Moreover, they need not even be literally essential to fail under Step Two. But they must be more than simply noteworthy. Courts must make a judgment as to whether the individual provision is of major significance to the statute, versus only minor significance. If the invalid provision is important to the entire statute or to a significant portion of the statute, then courts should infer that Congress would not have wanted the provisions connected with it to continue without it, and hold it either totally nonseverable or partially severable, respectively. As discussed below, Congress can heavily tilt the scale in favor of severability by including an express severability clause; however, absent such a clause, a provision is nonseverable if it is a prominent aspect of Congress's overall plan.

This has been implicit in severability doctrine since 1894, where the Court reasoned a statutory provision was severable because it was "simply in aid," as opposed to being an integral element "of the main purpose of the statute."[290] The Massachusetts high court explained why this is so several years after deciding *Warren*, when one provision "may have been the motive, inducement or consideration on which the other was founded, . . . they must stand or fall together."[291] Thus, the statutory scheme created by the statute would not be frustrated by the lack of the provision in question. The Justices unanimously restated this rule explicitly in 1987, holding that a court must discover "the original legislative bargain" codified in the statute,[292] and then determine whether the balance struck in

290. Reagan v. Farmers' Loan & Trust Co., 154 U.S. 362, 396 (1894).

291. Jones v. Robbins, 74 Mass. (8 Gray) 329, 339 (1857) (citing Warren v. Mayor of Charlestown, 68 Mass. (2 Gray) 84, 99 (1854)).

292. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987).

the statute is upset by the excision of the constitutionally-infirm provision.[293] If not, the abridged statute remains. But if that deal has become unbalanced, then the statute must fall.[294]

### B. *Various Types of Severability*

There are three possible outcomes to a severability inquiry whenever a provision is invalidated. The first is to strike down only the invalid provision. The second is to strike down all of the statute. And the third is an intermediate remedy to strike down more of the statute than just the invalid provision, but still retain a portion of the statute. Exploring these provisions at the theoretical level also requires the academy to answer an antecedent question that courts consider implicitly but rarely if ever expressly discuss: Of the words contained in a statute, precisely which words constitute a "provision?"

In answering these questions we must realize that statutes come in two forms. One option is that a statute is in a sense a grouping of legislative enactments—a series of short laws, sometimes related but other times not—that, for the sake of convenience, expediency, and avoiding redundancy are bundled together into a single bill. The other option is that an enacted statute represents a carefully balanced legislative deal, as shown in Part IV.A.2.b. It is much more likely that severance is possible under the first option, as the remaining provisions may be completely autonomous. It is much less likely that severance is possible under the second option, as the judge would have to explain why the original legislative bargain is not being undone by removing an invalid provision. (Even then, severability is possible when a court can show the invalid provision is incidental and insignificant to the legislative bargain.) As explained below, partial severability is sometimes possible with both types of statutes.

#### 1. Total Versus Partial Severability

When a court determines that an invalid provision cannot be severed, the court must then determine whether any part of the statute can be salvaged. If the unconstitutional provision can be

293. *See id.*

294. Various scholars have likewise seen this part of a severability inquiry as determining whether the invalid provision is essential to Congress's bargain. *See* Movsesian, *supra* note 27, at 62; *cf.* Walsh, *supra* note 6, at 770–71.

completely separated from the residuary of the statute, then it is said to be totally severable. Alternatively, if there are various other provisions that are so bound up with the invalid provision that they cannot be separated, while still other provisions are separable under the *Free Enterprise* test explained above, then the invalid provision is said to be partially severable. As the Eleventh Circuit explained it, "The fact that an invalid portion of a statute is not self-contained in separate sections does not prohibit the court from applying the severability rule to strike the invalid portion and to preserve the rest of the enactment."[295] Courts have more choices than a President does when evaluating legislation. To suggest that a court must either strike down only the invalid provision or strike down the entire statute is analogous to the choice a President faces of either signing an entire bill into law or vetoing the entire bill.[296] Courts are not so limited in their choices.

In fact, under modern doctrine, invalidating a statute *in toto* due to one invalid provision is the exception, not the rule. The most common remedy when a court invalidates a provision is to completely sever the unconstitutional provision from the rest of the statute, saving the rest of the Act. The Supreme Court has on occasion found statutes partially severable,[297] striking down provisions in addition to the unconstitutional provision, but still retaining much of the statute at issue. By contrast, the last significant case in which the Court held a provision completely nonseverable was when it invalidated an entire statute due to one faulty provision was in 1922—a case made all the more significant by the fact that the statute at issue included a severability clause.[298]

Often a court's severability analysis will be framed *ab initio* as a partial severability inquiry. This is seen when the court phrases the issue as whether the invalid provision can be severed from anything less than the entire statute. Consider this in formulaic terms, when X is the invalidated provision, Y is the portion of the statute containing the invalid provision, and Z is the entire statute. In such a scenario, X is a subset of Y, and Y in turn is a

295. Frazier v. Winn, 535 F.3d 1279, 1283 (11th Cir. 2008).

296. *See* Jona, *supra* note 197, at 712.

297. *E.g.*, R.R. Ret. Bd. v. Alton R.R. Co., 295 U.S. 330, 361 (1935) (holding several provisions of Railroad Retirement Act nonseverable).

298. Hill v. Wallace, 259 U.S. 44, 70–72 (1922).

subset of Z. Syllogistically, X is therefore a subset of Z. When a court asks whether X can be severed from Y, instead of whether X can be severed from Z, then the court's inquiry is a partial severability analysis. When a court thus defines its task as determining whether a provision is separable from a part of a statute and never even raises the question of total nonseverance, it implicitly holds *sub silentio* that at minimum a substantial part of the statute can be severed from the provision. In so doing, the court takes the possibility of total nonseverability entirely off the table.

In the seminal campaign finance case *Buckley v. Valeo,* the Court considered whether the two provisions at issue were severable from the remainder of Subtitle H of the Internal Revenue Code.[299] The Court held that the provisions were severable under the *Champlin* test,[300] only considering the question of whether the provisions were severable from the parts of the Federal Election Campaign Act (FECA) impacting Title 26 U.S.C., thus being a consideration of partial severability, without ever contemplating the question of total nonseverability which could potentially have also invalidated other provisions of FECA codified in Title 2 U.S.C.[301]

Even clearer, in *Planned Parenthood v. Danforth,* the severability inquiry was limited to whether one invalid provision within a single statutory section was severable from the remainder of that single statutory section.[302] The Court held it was not severable, concluding "that § 6(1) must stand or fall as a unit."[303] The Court invalidated that specific section, but never contemplated or discussed the possibility of nullifying the entire enactment.

One of the Supreme Court's watershed severability cases provides yet another example. In *Alaska Airlines* the Court was considering whether an invalid legislative-veto provision—which was one paragraph of a statute—could be severed from the Employee Protection Program which was created by section 43

299. 424 U.S. 1, 108 (1976).

300. *Id.* at 109. Recall that the *Champlin* test was the standard test for severability inquiries during this time, as *Buckley* predated *Alaska Airlines* and *Free Enterprise. See supra* Part II.B.

301. *See id.* at 108–09 (discussing whether the Federal Election Commission can exercise the powers conferred upon it). Title 26 U.S.C. is the Internal Revenue Code.

302. 428 U.S. 52, 83 (1976).

303. *Id.*

of a statute,[304] rather than the entire statute.[305] Justice Blackmun's opinion revolves around whether the statutory section at issue could function in the manner Congress intended without the invalid legislative-veto provision. Although, as already shown, *Alaska Airlines* specifically inquires whether the entire statute—not merely the statutory provision—can function as intended, a significant portion of the discussion nonetheless revolved around specifically how the implicated section would function without the invalid provision.

This suggests that the scope of severability inquiries—whether encompassing the whole statute or only part of a statute—is determined in *Free Enterprise* Step One by the provision's contextual setting in the court's opinion. That is to say, if part of the statute is rendered inoperable by the invalidation of one of that portion's provisions, but the remaining portions of the statute are still functional, then the provision is partially severable.

It is unclear how this approach translates to *Free Enterprise* Step Two. The Court's instructions are explicit to examine whether the *statute*—which would mean the whole of the statute—can function in the manner Congress intended. The Court never *explicitly* asks if the relevant *portion* of the challenged statute can function in its intended manner. It is not clear if this distinction in the case law means that answering the test's question in the negative must result in the statute's complete invalidation, or instead if it could result in excising only the infected part of a multi-part statute. At least one major authority from the early years of severability doctrine seemed to suggest the latter approach.[306] Which approach proves easier for a court to apply

304. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 683 (1987).

305. Justice Harry Blackmun begins his opinion for the Court by saying the Court was "consider[ing] whether [the] legislative-veto provision is severable from the remainder of the Act." *Id.* at 680. However, this language proves unhelpfully imprecise. That provision is found in section 43(f)(3) of the statute in question. *Id.* at 682. Justice Blackmun later clarifies that in *Alaska Airlines* the plaintiffs are making a constitutional challenge to the Employee Protection Program (EPP) that is created exclusively by section 43 of the statute, *see id.* at 680–82, and the Court is specifically affirming the judgment of the D.C. Circuit "that the legislative-veto clause is severable from the remainder of the EPP program." *Id.* at 683.

306. *See:*

> When neither the statute itself nor its legislative history afford any clue to the intention of the legislative body, the Court should look to the policy sought to be effectuated by the statute and decide whether that policy will be more nearly attained by partial application or by complete nullification of the law.

would likely depend on the specific facts of the statute in each case. If it is difficult to determine the intent underlying a statute as a whole, it stands to reason that it could sometimes be more difficult to determine whether a single section or subsection will function in the intended manner, but other times might be easier to assess the impact on only that implicated section.

Though there is an argument to the contrary. At first thought it is difficult to conjure a hypothetical statute in which excising a single provision results in the statute not functioning in Congress's intended manner, but going further to eradicate the entire section wherein the invalid provision is situated would restore that intended effect. While such an outcome seems unlikely, however, it is possible. You can imagine a section governing some aspect of the statute that Congress intends to play a secondary role, such as extraordinary powers that are only triggered under very unlikely circumstances as a contingency measure. Now imagine that the invalid provision is the provision containing the trigger, which is intended to severely limit the frequency with which these extraordinary powers are triggered. Under this hypothetical, to strike down the limiting provision would allow these reserve powers to be effectuated regularly or perhaps even be continuously in effect, rather than as powers that are rarely if ever invoked. A similar hypothetical would be if the invalid provision severely restricted the scope of extraordinary powers, but for which the President would wield tremendous authority. Under either hypothetical, removing one provision in the section would clearly upset Congress's "original legislative bargain" referenced in *Alaska Airlines*,[307] which is central to assessing severability under *Free Enterprise* Step Two. Under these facts, a court might be able to preserve Congress's intended functioning of the statute by holding the invalid provision partially severable, and invalidate the section in question but sever the entire section from the remainder of the statute. In terms of our earlier equation, it would be to say that X could not be severed from Y, but all of Y could be severed from Z. Courts will need to consider such hypotheticals in resolving this question, should such a case arise.

The idea of partial severability under myriad circumstances is consistent with the principles and rationale underlying

---

Stern, *supra* note 16, at 101.

307. 480 U.S. at 685.

severability doctrine examined in Part III, *supra.* A court assigns itself a limited role when it defines the judicial task as determining whether to invalidate only a single provision versus a small number of provisions (those surrounding the invalid provision, such as the remainder of a single section or even a single paragraph). It showcases severability as a doctrine of judicial restraint, leaving much of the public policy adopted by the people's elected representatives in force. This is consistent with the three *Ayotte* principles explored in Part III of minimalism, refraining from rewriting statutes, and vindicating congressional intent.[308]

But there can be no bright line rule, because under other circumstances partial severability would violate these principles. Courts must eschew partial severability whenever such an action—even when cloaked by an invocation of judicial restraint—essentially rewrites a statute or rebalances the statutory scheme embodied in the enactment. Severability remedies must be consistent with *all three* of the *Ayotte* principles; the first principle of minimalism cannot be utilized in such a way as to violate the principles of not rewriting a statute or subverting Congress's intent. This is reminiscent of the Court's reasoning (on a different severability matter) in *Free Enterprise,* where Chief Justice Roberts noted that the Court could excise a number of provisions of Sarbanes-Oxley as an alternative route to a constitutionalizing result, but rejected such a "blue-pencil" editorial approach as usurping Congress's role.[309] Whether a court strikes down only the invalid provision, versus part of the statute, versus all the statute, must be determined in consultation with not upsetting Congress's balancing of policy priorities, as such a rebalancing of a statute *de facto* rewrites the statute by substituting judges' policy preferences in derogation of Congress's prerogatives.

### a. *Defining "Provision"*

One question yet to be answered is how to define a statutory "provision." One prerequisite to determining whether to strike

308. *See* Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 328–30 (identifying the three principles that guide the remedy for constitutional flaws in statutes).

309. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3162 (2010).

down one or more provisions of a statute is to understand the term "provision" as the unit of analysis. This task is not as simple as some might guess, as case law reveals a broad diversity of combinations of words that together form a "provision" in any given statute.

"A text is clear if all or most persons, having the linguistic and cultural competence assumed by the authors of the text, would agree on its meaning."[310] The purpose of statutory text is to convey a particular meaning, declaring rights or imposing obligations and conditions, carrying the force of law. "If a message is unclear we ask the sender to repeat or amplify it until we no longer doubt what he meant to say."[311] This is true not only of statutory language, but indeed of all written language through which one person seeks to communicate to another.

Cases over the past forty years show the elasticity in length that can attend what the law denominates as a "provision." The Court in *Buckley* invalidated two provisions of the Federal Election Campaign Act of 1971.[312] The first of these, codified at 18 U.S.C. § 608(a) when *Buckley* was decided in 1976, consisted of nine sentence-length items, and the second, codified at that time at 18 U.S.C. § 608(c), consisted of eighteen sentence-length items.[313] That same year the Court in *Danforth* was considering whether the first sentence of section 6(1) of the Missouri statute at issue could be severed from the second sentence of section 6(1).[314] The Court in *Chadha* specifically found the one-chamber

310. Posner, *supra* note 210, at 187.

311. *Id.* at 188.

312. Buckley v. Valeo, 424 U.S. 1, 54, 58 (1976) (invalidating parts of Pub. L. No. 92-225, 86 Stat. 3 (1971), amended by Pub. L. No. 93-443, 88 Stat. 1263 (1974), codified in relevant part at 18 U.S.C. § 608(a), (c) (1970 Supp. IV)).

313. Space considerations make imprudent the full reproduction of those subsections here. Given that those invalidated provisions are no longer found in the current United States Code, subsection 608(a) can be found in the *Buckley* Court's appendix, *id.* at 187–89, and subsection 608(c) can be found in the same appendix, *id.* at 190–92.

314. Planned Parenthood v. Danforth, 428 U.S. 52, 83 (1976). The Missouri statute is reproduced in the Court's appendix. Although the severability inquiry concerned the first and second sentences of section 6(1), it contains a third sentence as well, and the entire section reads as follows:

> Section 6. (1) No person who performs or induces an abortion shall fail to exercise that degree of professional skill, care and diligence to preserve the life and health of the fetus which such person would be required to exercise in order to preserve the life and health of any fetus intended to be born and not aborted. Any physician or person assisting in the abortion who shall fail to take such measures to encourage or to sustain the life of the child, and the death of the child results, shall be deemed guilty of manslaughter and upon conviction

veto provision to be a "particular provision" of the statute.[315] That provision was section 244(c)(2) of the Immigration and Nationality Act.[316] So "provision" there referred to one subpart of one subsection of a statute, which was a full paragraph.[317] In *Regan v. Time* the Court invalidated an anti-counterfeiting provision then found at 18 U.S.C. § 504(1),[318] which is one statutory paragraph the length of multiple normal paragraphs,[319]

---

shall be punished as provided in section 559.140, RSMo. Further, such physician or other person shall be liable in an action for damages as provided in section 537.080, RSMo.

*Id.* at 85–86.

315. INS v. Chadha, 462 U.S. 919, 932 (1983).

316. *See id.* at 923, 932, 934 (referencing Immigration and Nationality Act of 1952, Pub. L. No. 414 § 244(c)(2), 66 Stat. 163, 214, codified at 8 U.S.C. § 1254(c)(2) (1980)).

317. The paragraph reads:

(2) In the case of an alien specified in paragraph (1) of subsection (a) of this subsection—

if during the session of the Congress at which a case is reported, or prior to the close of the session of the Congress next following the session at which a case is reported, either the Senate or the House of Representatives passes a resolution stating in substance that it does not favor the suspension of such deportation, the Attorney General shall thereupon deport such alien or authorize the alien's voluntary departure at his own expense under the order of deportation in the manner provided by law. If, within the time above specified, neither the Senate nor the House of Representatives shall pass such a resolution, the Attorney General shall cancel deportation proceedings.

Pub. L. No. 414 § 244(c)(2).

318. Regan v. Time, Inc., 468 U.S. 641, 659 (1984) (plurality opinion).

319. Section 504(1) allowed certain exceptions to general anti-counterfeiting requirements barring the reproduction of U.S. currency by providing:

Notwithstanding any other provision of this chapter, the following are permitted:

(1) the printing, publishing, or importation, or the making or importation of the necessary plates for such printing or publishing, of illustrations of—

(A) postage stamps of the United States,

(B) revenue stamps of the United States,

(C) any other obligation or security of the United States, and

(D) postage stamps, revenue stamps, notes, bonds, and any other obligation or other security of any foreign government, bank, or corporation for philatelic, numismatic, educational, historical, or newsworthy purposes in articles, books, journals, newspapers, or albums (but not for advertising purposes, except illustrations of stamps and paper money in philatelic or numismatic articles, books, journals, newspapers, or albums). Illustrations permitted by the foregoing provisions of this section shall be made in accordance with the following conditions—

(i) all illustrations shall be in black and white, except that illustrations of postage stamps issued by the United States or by any foreign government may be in color;

(ii) all illustrations (including illustrations of uncanceled postage stamps in color) shall be of a size less than three-fourths and more than one and one-half, in linear dimension, of each part of any matter so illustrated which is covered by subparagraph (A), (B), (C), or (D) of this paragraph, except that

and held this provision severable from the remainder of the Act.[320] In *Brockett* the Court invalidated part of a public morality statute involving sexually obscene material, and in so doing considered whether it would invalidate the entire statute.[321] In *Brockett*, the provision severed by the Court was a single clause of six words.[322] In *Alaska Airlines*, the Court severed a legislative veto in a statute deregulating airlines.[323] This provision was one paragraph of an otherwise-valid section of a statute.[324] In *New York*, the Court invalidated the take-title provision of a radioactive waste statute, a provision that was a lengthy sentence in one paragraph.[325] In *Booker*, the Court severed two separate

---

> black and white illustrations of postage and revenue stamps issued by the United States or by any foreign government and colored illustrations of canceled postage stamps issued by the United States may be in the exact linear dimension in which the stamps were issued; and
>
> (iii) the negatives and plates used in making illustrations shall be destroyed after their final use in accordance with this section.

*Id.* at 646 n.3 (quoting Act of Mar. 3, 1923, ch. 218, 42 Stat. 1437, as amended by Pub. L. No. 85-921, 72 Stat. 1771 (1958) (codified at 18 U.S.C. § 504(1) (1952))).

320. *Id.* at 653–54 (plurality opinion).

321. Brockett v. Spokane Arcades, Inc., 472 U.S. 491, 504–07 (1985).

322. *Id.* at 494 (quoting WASH. REV. CODE § 7.48A.010(8) (defining the term "prurient" as "that which incites lasciviousness or lust")).

323. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 697 (1987).

324. The legislative veto provision reads:

> (3) The Secretary shall not issue any rule or regulation as a final rule or regulation under this section until 30 legislative days after it has been submitted to the Committee on Commerce, Science, and Transportation of the Senate and the Committee on Public Works and Transportation of the House of Representatives. Any rule or regulation issued by the Secretary under this section as a final rule or regulation shall be submitted to the Congress and shall become effective 60 legislative days after the date of such submission, unless during that 60-day period either House adopts a resolution stating that that House disapproves such rules or regulations, except that such rules or regulations may become effective on the date, during such 60-day period, that a resolution has been adopted by both Houses stating that the Congress approves of them.

Airline Deregulation Act of 1978, Pub. L. No. 95-504 § 43(f)(3), 92 Stat. 1705, 1752 (codified at 49 U.S.C. App. § 1552(f)(3) (1982)). This was one paragraph in section 43 of the statute, found at 92 Stat. 1750–53.

325. New York v. United States, 505 U.S. 144, 174–77 (1992). The provision reads:

> If a State (or, where applicable, a compact region) in which low-level radioactive waste is generated is unable to provide for the disposal of all such waste generated within such State or compact region by January 1, 1996, each State in which such waste is generated, upon the request of the generator or owner of the waste shall take title to the waste, be obligated to take possession of the waste, and shall be liable for all damages directly or indirectly incurred by such generator or owner as a consequence of the failure of the State to take possession of the waste as soon after January 1, 1996, as the generator or owner notifies the State that the waste is available for shipment.

provisions of the Sentencing Reform Act, both of which were subsections, one of which is a single paragraph and the other which can be characterized as either one or two paragraphs.[326] In *Ayotte*, the challenged aspect of the statute was that it did not

---

Low Level Radioactive Waste Policy Act, Amended, Pub. L. No. 99-240 § 5(d)(2)(C), 99 Stat. 1842, 1851 (1986) (codified at 42 U.S.C. § 2021e(d)(2)(C) (1982)).

326. United States v. Booker, 543 U.S. 220, 270–71 (2005). The first provision invalidated by the Court read:

> Application of guidelines in imposing sentence.—(1) In general.—Except as provided in paragraph (2), the court shall impose a sentence of the kind, and within the range, referred to in subsection (a)(4) unless the court finds that there exists an aggravating or mitigating circumstance of a kind, or to a degree, not adequately taken into consideration by the Sentencing Commission in formulating the guidelines that should result in a sentence different from that described. In determining whether a circumstance was adequately taken into consideration, the court shall consider only the sentencing guidelines, policy statements, and official commentary of the Sentencing Commission. In the absence of an applicable sentencing guideline, the court shall impose an appropriate sentence, having due regard for the relationship of the sentence imposed to sentences prescribed by guidelines applicable to similar offenses and offenders, and to the applicable policy statements of the Sentencing Commission.

Sentencing Reform Act of 1984, Pub. L. No. 98-473, 98 Stat. 1987 (1984) (codified at 18 U.S.C. § 3553(b)(1) (2000 Supp. IV)). The second provision read:

> Consideration.—Upon review of the record, the court of appeals shall determine whether the sentence—
>
> (1) was imposed in violation of law;
>
> (2) was imposed as a result of an incorrect application of the sentencing guidelines;
>
> (3) is outside the applicable guideline range, and
>
> (A) the district court failed to provide the written statement of reasons required by section 3553(c);
>
> (B) the sentence departs from the applicable guideline range based on a factor that—
>
> (i) does not advance the objectives set forth in section 3553(a)(2); or
>
> (ii) is not authorized under section 3553(b); or
>
> (iii) is not justified by the facts of the case; or
>
> (C) the sentence departs to an unreasonable degree from the applicable guidelines range, having regard for the factors to be considered in imposing a sentence, as set forth in section 3553(a) of this title and the reasons for the imposition of the particular sentence, as stated by the district court pursuant to the provisions of section 3553(c); or
>
> (4) was imposed for an offense for which there is no applicable sentencing guideline and is plainly unreasonable.
>
> The court of appeals shall give due regard to the opportunity of the district court to judge the credibility of the witnesses, and shall accept the findings of fact the district court unless they are clearly erroneous and, except with respect to determinations under subsection (3)(A) or (3)(B), shall give due deference to the district court's application of the guidelines to the facts. With respect to determinations under subsection (3)(A) or (3)(B), the court of appeals shall review de novo the district court's application of the guidelines to the facts.

*Id.* (codified at 18 U.S.C. § 3742(e) (2000 and Supp. IV)).

provide an exception to parental notification of a minor's abortion in the event of a medical emergency.[327] Thus, this challenge of a *lack* of a provision was essentially a challenge to the statute's existing medical provision, arguing that this medical provision was not broad enough to cover emergencies that were not deadly. This provision was a single compound sentence.[328] Most recently, in *Free Enterprise* the Court considered whether invalidating the tenure provision of the Sarbanes-Oxley Act was unconstitutional, wherein the challenged provision was actually two different clauses of the statute.[329] Thus, a "provision" can range in length from several words the length of a single simple clause to oversized paragraphs exceeding a page in size. A provision can be in the form of one group of words, or disjointed groups of words, or even the lack of certain words. And a single provision in these examples from recent cases can range anywhere from six words to 281.[330]

---

327. Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 324 (2006).

328. *See id.* This provision allowed for an abortion if the provider deemed the minor's life in danger. The provision states as one exception to notifying a parent of an impending abortion: "The attending abortion provider certifies in the pregnant minor's medical record that the abortion is necessary to prevent the minor's death and there is insufficient time to provide the required notice." N.H. REV. STAT. ANN. § 132:26(I)(a) (Supp. 2004) (repealed 2007).

329. *See* Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3148, 3161 (2010) Of these two provisions, the primary invalidity is found in provision that allows for Board members to be removed, "in accordance with [§ 7217(d)(3)], for good cause shown." 15 U.S.C. § 7211(e)(6) (2010). In invalidating these seven words, the Court also invalidated the provision incorporated by reference, which reads that a member can be removed once the Securities and Exchange commission finds, "on the record" and "after notice and opportunity for a hearing," that a Board member:

> (A) has willfully violated any provision of this Act, the rules of the Board, or the securities laws;
>
> (B) has willfully abused the authority of that member, or
>
> (C) without reasonable justification or excuse, has failed to enforce compliance with any such provision or rule, or any professional standard by any registered public accounting firm or any associated person thereof.

*Id.* § 7217(d)(3).

330. *Compare* provision in *supra* note 322 (six words in length), *with* provision in *supra* note 319 (281 words in length). Also the word counts from the two invalidated provisions in *Booker* are 271 and 148, with a combined total of 419 words. *See supra* note 326. Another case during this time frame briefly discussing severability is *Randall v. Sorrell*, 548 U.S. 230 (2006). However, the Court's application of severability in *Randall* did not make entirely clear which provisions it found valid versus invalid, summarily holding that some of the campaign contribution limit provisions which might be constitutional in the Vermont campaign-finance statute at issue could not be severed from those limits that were unconstitutional, invalidating them all. *Id.* at 262. Thus it is not easy or necessarily useful to provide an exact word count on the invalid provisions at issue in that case.

Not only that, but *Randall* was evaluating a Vermont statute. When federal courts consider state statutes, severability is a question of state law. *See* Leavitt v. Jane L., 518

### b. *Definitional Challenge Endemic to All Human Communication in Reducing Thoughts into Words*

Although this topic deserves extensive discussion beyond the scope of this Article, it should be noted that the problem in defining "provision" is not unique to severability doctrine, but rather endemic to all written communication.

The problems attending communication are more pronounced with the written word. Writings are often one-way communications, in which the writer is "speaking" and the reader is "listening," and for which the reader has no ability to respond in a way that the writer can immediately perceive.[331]

Whether spoken or written, the value of the use of language is measured by the effectiveness in which ideas are being successfully perceived and comprehended by the intended audience. The most eloquent speech or treatise on the most important issue is of no value if the intended audience is unable to understand the message being conveyed.

Thus the purpose of language is to reduce human thought to words.[332] However, completely effective and efficient communication is impossible.[333] There are always slight nuances of intended meaning that are not picked up in language, either

---

U.S. 137, 139 (1996). So this case is not especially helpful in understanding federal severability doctrine, and should not be invoked in evaluating federal statutes.

331. The written word is typically a monologue, not a dialogue. This fact is one of the reasons clarity is of paramount importance in writing, since there is no opportunity for the writer to realize that his audience does not understand what is being conveyed, and thus to attempt to remedy the situation via restatement or analogy.

332. Those words are then transmitted to a recipient via a medium accessible by one of the recipient's five senses, either through audible communication—typically speech—or visual communication, such writing, and sometimes images. (Alternatively, transmission of ideas can be through a kinesthetic or tactile sense, such as Braille for the blind. It is also theoretically possible that communication could be achieved through the olfactory or gustatory senses (smell and taste, respectively), though I am unfamiliar with any successful attempts to do so with a system sufficiently sophisticated to constitute a "language" that can communicate complex ideas.) If the recipient is able to receive a message through being conscious and aware, and able to decipher the message by an adequate mastery of the language in which the message is codified, then to the extent that the recipient accurately processes the sender's thoughts, communication has been achieved.

333. No series of words can encapsulate a human thought in perfect detail. Whenever thought is reduced to words—either spoken or written—there is always some variance between the thought and the words, because words themselves are merely a human construct; every language is a cumulative attempt by a collection of fallible human beings to devise a medium of communication, designed to approximate as closely as possible an accurate encapsulation of human thought. But even the slightest misuse of a word—either because the speaker is misusing it, or because the audience does not correctly understand its definition—results in some degradation of the message en route from origin to final destination.

because no word exists that perfectly portrays the concept in the speaker's mind, or because the speaker's primary thought is accompanied by a relevant secondary or even tertiary thought for which the speaker does not attempt to find accompanying words to add to the message sent to the listener.

The application here is that one of the fundamental challenges with severability inquiries is that a provision is a legislative "thought" or "idea" reduced to writing. Each statutory "provision" is thus a "thought" of Congress communicated through the written word. What is the relationship between thoughts? Where does one "thought" end and a completely separate "thought" begin? Such a termination of one thought and initiation of the following thought could be after the next comma, or semicolon, or period, or paragraph break, or section break, or chapter break. Such breakpoints can be difficult to demarcate. Or more complicated yet, as just discussed regarding *Free Enterprise*, the provision at issue was actually contained in two separate, nonconsecutive sections of the Sarbanes-Oxley Act, demonstrating that a single thought can be embodied in part in a disjointed fashion through nonconsecutive statutory texts.

This is the quandary courts face with severability inquiries. It is determining at what level of specificity to define a "provision," and then determining the proximity in thought, purpose, and intent, between an invalid provision and the other provisions in the statute. Whether an invalid provision can be severed either from all of the remaining statute or part of the remaining statute is predicated on this conceptual exercise.

### 2. Multiprovisional Severability

Another issue that the Supreme Court has not directly confronted is multiprovisional severability. There may be instances in which a statute could survive the invalidation of one provision that is not so important to the statutory scheme so as to justify erasing the legislation, but in which a second and perhaps a third provision is also invalid. Even if none of these provisions would be fatal to the statute, invalidating multiple provisions can reach a tipping point after which a court must invalidate *in toto* to prompt Congress to reconsider how to address the policies impacted by the statute.

Multiprovisional severability can arise either in a partial severability context or in a total severability context. In the

former, a court examines the impact of multiple invalid provisions on a single part of a statute. In the latter, a court examines the impact on the entire statute.

Visualize the unit of analysis—whether the entire statute or part of a statute such as a chapter—as the hull of a ship. Ships have watertight hatches and bulkheads to contain damage in the event of a hull breach.[334] According to experts, amongst other causes, the famous *Titanic* sank in the Atlantic Ocean not because of a single massive hull breach, but rather because of a series of smaller breaches in different compartments of the ship.[335] While the *Titanic* could have survived one such breach, the aggregate impact of multiple breaches is what doomed the storied ship.[336] One can imagine statutes with two invalid provisions, where the statute can still function in the manner Congress intended without any one of those provisions. Perhaps one of the provisions can partially achieve the effects primarily expected to arise from the other. However, without both provisions, the functioning of the statute would be so impaired that the entire statute becomes unsalvageable. By punching enough holes in the hull, eventually the *Titanic* will sink.

Partial severability can also become self-perpetuating, such that it becomes a form of multiprovisional severability. With each additional provision a court finds nonseverable from the invalid provision, the judge must then ask whether these additional provisions are fatally wounded by the implicated provision, in addition to the inherently-invalid provision. Begin with a scenario where Section A is invalid. Section B is so intertwined with Section A as to be nonseverable, hence Section B is also invalidated. Section C is not directly dependent on Section A, and so at first blush is severed. But while Section C can still function without Section A, it can no longer function in the manner Congress intended without Section B. As a result, Section C must also be struck down, despite its lack of functional proximity to Section A. So invalidity radiates from the unconstitutional provision, not only carrying down provisions sufficiently impacted by invalidating the unconstitutional

334. *See* THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE 251 (3d ed. 1996).

335. THE ENCYCLOPEDIA BRITANNICA, "Navy Architecture: Buoyancy," in *Titanic*, *available at* http://www.britannica.com/titanic/buoyancy.html (last visited Nov. 9, 2011).

336. *See id.*

provision, but also carrying down provisions not directly tied to the infirm provision.

## V. SEVERABILITY CLAUSE DICTATES WHETHER PRESUMPTION OF SEVERABILITY EXISTS

Many statutes—like many contracts—contain a severability clause. Typically, such a provision reads, "If any provision of this statute is found invalid or otherwise unenforceable, the remainder shall continue in full force in effect," or words of the nature. A severability clause—alternatively designated as a severance clause, separability clause, or saving clause—has two components regarding the overall written document, whether the document is an enrolled statute or some form of contractual instrument. The first of those clauses is a condition precedent, announcing the clause as a contingency provision in the event that one or more of the other provisions in the document cannot be given full or partial effect. The second is the operative clause, providing that notwithstanding the faulty provision, the remainder of the document should still be effectual. Such a severability clause is a clear statement of congressional intent, just as a severability clause in a contract is a clear statement of the parties' intent.

Severability clauses have provided varying roles over the decades. They became common around the turn of the twentieth century and into the 1930s in an effort to counteract courts' growing tendency to strike down statutes entirely on account of a single invalid provision.[337] What one scholar labels *dicta* in Supreme Court cases circa 1900 suggested a presumption against severability when a statutory provision was found invalid,[338] consistent with both lesser judicial and scholarly authorities.[339]

337. *See* Nagle, *supra* note 5, at 218.

338. *Id.* at 218 (citing El Paso & Ne. Ry. Co. v. Gutierrez, 215 U.S. 87, 97 (1909); Employers' Liability Cases, 207 U.S. 463, 501 (1908)). It is doubtful that these statements are *dicta*. Even if they are, however, the Supreme Court's statements by the 1930s were clearly part of the Court's holding, and no mere *dicta*. *See, e.g.*, Utah Power & Light Co. v. Pfost, 286 U.S. 165, 184–85 (1932).

339. *Id.* at 218 n.79 (citing Skagit Cnty. v. Stiles, 39 P. 116, 116 (Wash. 1894); THOMAS M. COOLEY, TREATISE ON THE CONSTITUTIONAL LIMITATIONS WHICH REST UPON THE LEGISLATIVE POWER OF THE STATES OF THE AMERICAN UNION 248 n.1 (Victor H. Lane ed., 7th ed. 1903)).

Whether or not it was ever *dicta*, the Supreme Court reaffirmed this presumption of nonseverability as part of the Court's holding in 1935,[340] and then added in 1936:

> In the absence of such a provision, the presumption is that the Legislature intends an act to be effective as an entirety—that is to say, the rule is against the mutilation of a statute; and if any provision be unconstitutional, the presumption is that the remaining provisions fall with it. The effect of the statute is to reverse this presumption in favor of inseparability and create the opposite one of separability. Under the nonstatutory rule, the burden is upon the supporter of the legislation to show the separability of the provisions involved. Under the statutory rule, the burden is shifted to the assailant to show their inseparability. But under either rule, the determination, in the end, is reached by applying the same test—namely, What is the intent of the law-makers?[341]

With so much riding on the presence or absence of a severability clause, legislatures had a strong incentive to include such a clause if severability was intended.

The impetus behind careful legislative drafting increased once the Supreme Court clearly held in 1932 and again in 1936 that severability clauses created a presumption of severability for statutes containing one.[342] If a severability clause is included, "when validity is in question, divisibility and not integration is the guiding principle. Invalid parts are to be excised and the remainder enforced. When we are seeking to ascertain the congressional purpose, we must give heed to this explicit declaration."[343]

While still prudent, the advisability of severability clauses became less important when the Supreme Court held in *Alaska Airlines* that the lack of a severability clause did not result in a

340. R.R. Ret. Bd. v. Alton R. Co., 295 U.S. 330, 361–62 (1935).

341. Carter v. Carter Coal Co., 298 U.S. 238, 312 (1936); *accord* Utah Power & Light Co. v. Pfost, 286 U.S. 165, 184–85 (1932).

342. *Carter*, 298 U.S. at 312; *see also* Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 235 (1932) (noting that the invalidity of a portion of an act does not render the entire act invalid, unless it is clear that the legislature would not have enacted the legislation without the invalid portion).

343. Elec. Bond & Share Co. v. Sec. & Exch. Comm'n, 303 U.S. 419, 434 (1938). This principle endured until as recently as 1983, as can be seen from the *INS* Court's holding that an unambiguous severability clause "gives rise to a presumption that Congress did not intend the validity of the Act as a whole, or of any part of the Act, to depend upon whether the veto clause of § 244(c)(2) was invalid." INS v. Chadha, 462 U.S. 919, 932 (1983).

presumption against severability,[344] overruling the principle discussed above that had governed in the late nineteenth and early twentieth century. Since 1987 it has been indisputable that no statute is presumed nonseverable, regardless of whether there is a severability clause in the challenged statute. Nevertheless, lawyers are well-advised not to rely on the absence of a contrary presumption or even on favorable extraneous sources such as legislative history, since the general rule in statutory construction for ascertaining legislative intent is that "the authoritative statement is the statutory text, not the legislative history . . . ."[345]

A severability clause is not a panacea which neutralizes every attempt to nullify an entire statute. One authority argues that the presence of a severability clause should always be conclusive in favor of severability,[346] but the courts refuse to adopt that theory. Severability clauses in statutes are an aid in construction,[347] as previously noted similar to such clauses' effects in contracts.[348] "Whether the provisions of a statute are so interwoven that one being held invalid the others must fall, presents a question of statutory construction and of legislative intent, to the determination of which the statutory provision becomes an aid."[349]

344. *Alaska Airlines*, 480 U.S. at 686. This idea had been evolving for some time. The D.C. Circuit held that *Chadha* was the case in which the Supreme Court created a presumption of severability both when the truncated statute remains operative and also when a severability clause was present. *See* City of New Haven v. United States, 809 F.2d 900, 905 & n.15 (D.C. Cir. 1986). While Judge Douglas Ginsburg in his opinion for the circuit court presaged the Supreme Court's *Alaska Airlines* decision the following year, it would be more accurate to state that *Champlin* first held a severability clause creates a presumption of severability, and that *Chadha* strongly suggested this presumption should apply even without a severability clause. But it was not until *Alaska Airlines* that the Supreme Court elevated this suggestion to a holding.

345. Exxon Mobile Corp. v. Allapattah Servs., Inc., 545 U.S. 546, 568 (2005); *see also* Hoffman Plastic Compounds, Inc. v. NLRB, 535 U.S. 137, 149–50 n.4 (2002). Those considering such extratextual sources must also keep in mind that the Supreme Court has held that "extrinsic aids to construction may be used to solve, but not to create, an ambiguity." Chamber of Commerce of U.S.A. v. Whiting, 131 S. Ct. 1968, 1980 (2011) (quoting United States v. Shreveport Grain & Elevator Co., 287 U.S. 77, 83 (1932)) (internal citation marks omitted).

346. *See* Schumsky, *supra* note 183, at 246–52 (arguing that severability clauses do not infringe upon the Article III powers of judges).

347. *See* Dorchy v. Kansas, 264 U.S. 286, 290 (1924) (dictum).

348. *E.g.*, Eckles v. Sharman, 548 F.2d 905, 909 (10th Cir. 1977) (quoting Moffat Tunnel Improvement Dist. v. Denver & S.L. Ry., 45 F.2d 715, 731 (10th Cir. 1930)); Zerbetz v. Alaska Energy Ctr., 708 P.2d 1270, 1283 (Alaska 1985); *see also* RESTATEMENT *supra* note 283, at § 184 cmt. a; Movsesian, *supra* note 27, at 46–56.

349. Carter v. Carter Coal Co., 298 U.S. 238, 313 (1936).

But while severability clauses are rightly described as an aid, rather than a dispositive declaration, they are extremely valuable if a legislature truly prefers part of its new statute to no statute at all. Under modern doctrine, a severability clause establishes a strong presumption of severability.[350] And the strength of this presumption is sufficient to save most of the provisions of a statute in the vast majority of cases wherein a single provision is found invalid.

Although no federal court has yet had occasion to either follow or set at naught a nonseverability clause, on rare occasions state courts have severed a provision when the state legislature included a nonseverability clause.[351] While sovereign states are free to develop their own theory of severability to govern judicial review of their own laws, this should never happen in federal courts concerning federal statutes, as it is utterly indisputable that Congress's fundamental intent for the statute is subverted by retaining the residual statute when Congress expressly declares that it does not wish the statute to endure if part of it is removed.

### A. *No Presumption of Severability Without a Severability Clause*

One area of severability where confusion seems especially significant centers on whether statutes are presumed severable. Presumptions allow one litigant to prove one fact and presume another, shifting the burden of production to the other party.[352] A presumption is where, as a matter of judicial policy, a court infers something to be true because there is some likelihood that the presumed fact is true.[353] A presumption occurs when "finding [a] predicate fact . . . produces a required conclusion in the absence of explanation."[354] Presumptions discussed in the

350. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 686 (1987). This rule from *Alaska Airlines* is consistent with decades of previous severability precedents. *See* Elec. Bond & Share Co. v. Sec. & Exch. Comm'n, 303 U.S. 419, 434 (1938); Williams v. Standard Oil Co., 278 U.S. 235, 242 (1929); Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 235 (1932); Utah Power & Light Co. v. Pfost, 286 U.S. 165, 184 (1932).

351. Israel E. Friedman, Comment, *Inseverability Clauses in Statutes*, 64 U. CHI. L. REV. 903, 907–08 (1997).

352. David F. Johnson, *The Use of Presumptions in Summary Judgment Procedure in Texas and Federal Courts*, 54 BAYLOR L. REV. 605, 605 (2002).

353. 9 JOHN H. WIGMORE, EVIDENCE § 2491 (3d ed. 1940). Some scholars believe that the benefits theoretically conferred by presumptions are often outweighed by the confusion or uncertainty such presumptions create. Edmund M. Morgan, *Presumptions*, 12 WASH. L. REV. 255, 280 (1937). But that topic is beyond the scope of this Article.

354. St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 506 (1993).

context of litigation are specifically rebuttable presumptions,[355] since irrebuttable presumptions are *per se* true as a legal matter because by definition there is nothing the opposing side can say to overcome the court's presumption.

Although presumptions often exist for questions of fact (hence the prevalence of presumptions in evidentiary issues in trial-level proceedings),[356] there are also presumptions in questions of law. Congress is "free to adopt presumptions for policy reasons."[357] Courts do the same. While presumptions are often facts presumed true as evidentiary matters, the same logic governs court presumptions on rules of law.

Such presumptions are ubiquitous in judicial review.[358] For example, statutes burdening fundamental rights are presumed invalid.[359] The same is true for statutes discriminating on the basis of suspect classifications.[360] Apparently statutes differentiating on the basis of quasi-suspect classifications also trigger a presumption of invalidity, as the burden seems to shift to the government to defend the statute's constitutionality under even the intermediate form of heightened scrutiny.[361]

355. "A rebuttable presumption is a rule of evidence under which, once a basic fact or a group of basic facts (Facts A) have been established, the fact finder also must accept the presumed fact (Fact B) that follows from the basic fact unless the presumption is rebutted." Yen P. Hoang, Note, *Assessing Environmental Damages After Oil Spill Disasters: How Courts Should Construe the Rebuttable Presumption Under the Oil Pollution Act,* 96 CORNELL L. REV. 1469, 1480 (2011) (citing 1 JACK B. WEINSTEIN & MARGARET A. BERGER, WEINSTEIN'S FEDERAL EVIDENCE § 301.02[1] (2d ed. 2000)).

356. *See, e.g.*, JAMES B. THAYER, A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW 336–43 (Boston, Little, Brown & Co. 1898); David W. Louisell, *Construing Rule 301: Instructing the Jury on Presumptions in Civil Actions and Proceedings*, 63 VA. L. REV. 281, *passim* (1977); Morgan, *supra* note 353, at *passim.*

357. Gen. Elec. Co. v. U.S. Dep't of Commerce, 128 F.3d 767, 772 (D.C. Cir. 1997) (quoting Chem. Mfrs. Ass'n v. Dep't of Transp., 105 F.3d 702, 705 (D.C. Cir. 1997)).

358. State courts are presumed to have concurrent jurisdiction with federal courts over federal claims, absent Congress expressing a contrary intent. Tafflin v. Levitt, 493 U.S. 455, 459 (1990). Federal statutes are presumed not to have extraterritorial reach unless Congress manifests a clear intent to the contrary. Sale v. Haitian Ctrs. Council, 509 U.S. 155, 173 (1993) (citations omitted). Words identically used in multiple parts of the same statute are generally presumed to have the same meaning throughout. IBP, Inc. v. Alvarez, 546 U.S. 21, 34 (2005).

359. *E.g.*, United States v. Am. Library Ass'n, 539 U.S. 194, 235 (2003) (holding burdens on free speech are presumed invalid).

360. *E.g.*, Gratz v. Bollinger, 539 U.S. 244, 270 (2003) (applying strict scrutiny to a racial-preference program).

361. *See, e.g.*, Heckler v. Mathews, 465 U.S. 728, 744 (1984) ("[T]he party seeking to uphold a statute that classifies individuals on the basis of their gender must carry the burden of showing . . . that the classification serves important governmental objectives and that the discriminatory means employed are substantially related to the achievement of those objectives.") (quoting Miss. Univ. for Women v. Hogan, 458 U.S. 718, 724–25

Others laws—such as those subject to rational-basis review—carry a presumption of validity.[362] Such presumptions form a significant part of federal courts' function in performing judicial review.

Invalidating individually-valid provisions because they are so connected to an invalid provision—sometimes even in the face of a severability clause expressing Congress's intent—is another aspect of courts' power of judicial review.[363] Although some authors suggest statutes are presumed severable with or without a severability clause, the strength of that presumption decreases significantly when Congress does not write an express clause into the enactment. The Supreme Court has also not given any detailed instruction on what the standard of proof is when a clause is present or when it is absent; all that can be definitively asserted is that, when the presumption favors severability, the burden falls on the challenger to make the case for nonseverability to overcome the presumption.

But the Supreme Court has never even adopted such a presumption when a statute lacks a severability clause. Various scholars claim a presumption of implied severability exists.[364] But as Professor John Nagle observes in his seminal Article in *North Carolina Law Review*, the High Court has declined to adopt such a presumption.[365] Four Justices applied a presumption of severability in *Regan v. Time*.[366] But none of the other Justices in that case supported this position, so it is not controlling authority under the *Marks* rule.[367] Nonetheless, at least four federal circuit courts subsequently adopted this presumption.[368]

---

(1982)) (internal quotation marks omitted). Such gender-based statutes are subject to intermediate scrutiny in judicial review. Clark v. Jeter, 486 U.S. 456, 461 (1988).

362. Williamson v. Lee Optical of Okla., Inc., 348 U.S. 483, 488 (1955); *see also* United States v. Carolene Prods. Co., 304 U.S. 144, 152 (1938).

363. *See* Patrick O. Gudridge, Marbury *at 200: Judicial Supremacy Today: The Office of the Oath*, 20 CONST. COMMENT. 387, 401–02 (2003).

364. *See, e.g.*, Jona, *supra* note 197, at 704–05; Schumsky, *supra* note 183, at 243.

365. Nagle, *supra* note 5, at 220–21.

366. 468 U.S. 641, 652–53 (1984) (plurality opinion).

367. *See* Marks v. United States, 430 U.S. 188, 193 (1977) ("When a fragmented Court decides a case and no single rationale explaining the result enjoys the assent of five Justices, the holding of the Court may be viewed as that position taken by those Members who concurred in the judgments on the narrowest grounds.") (internal quotation marks omitted).

368. Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health and Human Servs., 648 F.3d 1235, 1321 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400); United States v. Ameline, 376 F.3d 967, 981 (9th Cir. 2004); Nat'l Treas. Emp. Union v. United States,

Although the Supreme Court has issued no clear holding on whether such a presumption exists, the Court has suggested it does not. The Court said nothing to elevate the *Regan* plurality's position to a holding three years later in *Alaska Airlines* or its subsequent severability case law. Instead, the *Alaska Airlines* Court held "that the inclusion of [a severability] clause *creates* a presumption that Congress did not intend the validity of the statute in question to depend on the validity of the constitutionally offensive provision."[369] The word "create" means to "bring into being" or "to give rise to" or "produce."[370] Thus to say a severability clause creates a presumption of severability implicitly presupposes that no such presumption exists without a clause. To bring something into being presupposes it was not already in being. To produce something presupposes that it did not exist before it was produced.

This is also the most logical position that reconciles modern severability doctrine with early severability doctrine. As previously discussed, in the first five decades of the Supreme Court's severability jurisprudence—an era in which severability clauses were rare—the Court routinely struck down entire statutes due to a single invalid provision.[371] Once severability clauses became common the typical result was only to strike down the invalid provision and salvage the remaining statute.[372] These later cases illustrate the judiciary giving effect when possible to Congress's intent as expressed in the text of the statute at issue in each case.

Without such a clause, courts must resort to extrinsic factors to determine legislative intent and are free to invalidate all of a statute when the court concludes there is significant doubt as to whether Congress would have enacted the remaining statute. The most recent restatement of severability doctrine is fully consistent with this approach; when the *Free Enterprise* Court severed the tenure-restriction provisions of the Board members in Sarbanes-Oxley it found that "*nothing* in the statute's text or historical context makes it evident that Congress, faced with the

990 F.2d 1271, 1278 (D.C. Cir. 1993); Cmty. For Creative Non-Violence v. Turner, 893 F.2d 1387, 1394 (D.C. Cir. 1990); News Am. Pub., Inc. v. FCC, 844 F.2d 800, 802 n.1 (D.C. Cir. 1988); Doyle v. Suffolk Cnty., 786 F.2d 523, 528 (2d Cir. 1986).

369. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 686 (1987) (emphasis added).

370. AMERICAN HERITAGE DICTIONARY, *supra* note 334, at 438.

371. *See, e.g.*, cases cited *supra* notes 56–64, 76–84 and accompanying text.

372. *See, e.g.*, cases cited *supra* notes 87–93 and accompanying text.

limitations imposed by the Constitution, would have preferred no Board at all to a Board whose members are removable at will."[373] Finding absolutely no intrinsic or extrinsic evidence of Congress's intent that the statute should fall due to the tenure protections on Board members, the Court concluded the unconstitutional part severable.

When a statute includes a severability clause, the language by which the Supreme Court describes the showing that must be made to overcome the presumption of severability looks like clear and convincing evidence. As the Court held in *Carter*:

> Under the statutory rule, the presumption must be overcome by the considerations which establish 'the clear probability that the invalid part being eliminated the Legislature would not have been satisfied with what remains,' or . . . 'the clear probability that the Legislature would not have been satisfied with the statute unless it had included the invalid part.'[374]

By invoking the term "*clear* probability"—as opposed to "probability" without an augmenting adjective—not just once, but twice, the Court is signaling a standard well beyond a mere preponderance of the evidence.

Case law elaborates on one other aspect of the burden challengers must satisfy to defeat severability. The Supreme Court described the effect of a severability clause as one that:

> furnishes assurances to courts that they may properly sustain separate sections or provisions of a partly invalid act without hesitation or doubt as to whether they would have been adopted, even if the legislature had been advised of the invalidity of part. But it does not give the court power to amend the act.[375]

Prior to *Alaska Airlines*, the D.C. Circuit held that the presumption is overcome only by "strong evidence" that Congress would not have enacted the statute if it knew that the invalid provision could not be a component of the legislation.[376] Then in *Alaska Airlines* the Supreme Court held that when

373. Free Enter. Fund v. Pub. Co. Accounting Oversight Bd., 130 S. Ct. 3138, 3162 (2010) (emphasis added) (internal quotation marks omitted).

374. Carter v. Carter Coal Co., 298 U.S. 238, 312–13 (1936) (quoting Williams v. Standard Oil Co., 278 U.S. 235, 241 (1929) and Utah Power & Light Co. v. Pfost, 286 U.S. 165, 184–85 (1932)).

375. Hill v. Wallace, 259 U.S. 44, 71 (1922).

376. City of New Haven v. United States, 809 F.2d 900, 905 (D.C. Cir. 1987).

considering the burden to be met to overcome a presumption of severability:

> The inquiry is eased when Congress has explicitly provided for severance by including a severability clause in the statute. This Court has held that the inclusion of such a clause creates a presumption that Congress did not intend the validity of the statute in question to depend on the validity of the constitutionally offensive provision. In such a case, unless there is strong evidence that Congress intended otherwise, the objectionable provision can be excised from the remainder of the statute. In the absence of a severability clause, however, Congress' silence is just that—silence—and does not raise a presumption against severability.[377]

This is because a severability clause "serves to assure the courts that separate sections or provisions of a partly invalid act may be properly sustained 'without hesitation or doubt as to whether they would have been adopted, even if the Legislature had been advised of the invalidity of part.'"[378]

As discussed above, even in the absence of a severability clause, some lower courts will apply a presumption of implied severability.[379] But since severability under *Free Enterprise* Step

---

377. Alaska Airlines Inc., v. Brock, 480 U.S. 678, 686 (1987).

378. *Williams*, 278 U.S. at 241 (quoting Hill v. Wallace, 259 U.S. 44, 71 (1922)).

379. It should also be noted that there is a circuit split on whether a court should factor removal of a severability clause during legislative debate into a severability analysis. Very rarely, there are statutes for which early versions of the bill contain a severability clause, but the legislature removes the clause before final passage. The Ninth Circuit briefly held on another matter:

> When Congress deliberately makes a decision to omit a particular provision from a statute—a decision that it is aware may well result in the statute's wholesale invalidation . . . we would not be faithful to its legislative intent were we to devise a remedy that in effect inserts the provision into the statute contrary to its wishes. Such an action would be inconsistent with our proper judicial role.

Planned Parenthood Fed'n of Am. v. Gonzales, 435 F.3d 1163, 1187 (9th Cir. 2006), *rev'd sub nom. on other grounds* Gonzales v. Carhart, 550 U.S. 124 (2007). This analogy is not precise, in that there the circuit court was considering an abortion statute where Congress expressly declined to include a health exception to an abortion restriction, when Supreme Court precedent clearly required such a provision. But aspects of the Ninth Circuit's reasoning is analogous to Congress opting not to include a severability clause in a statute that Congress suspects will be challenged in court.

But on that same issue, the Eleventh Circuit noted that the district court in the case before it considered removal of the clause "strong evidence that Congress recognized the Act could not operate as intended without the" invalid provision, but concluded that the "district court pushes this inference too far." Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health and Human Servs., 648 F.3d 1235, 1322 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400). Although the reference to *strong* evidence in such a statement in isolation

Two turns on congressional intent, the question becomes determining not only to what degree the presumption of severability is diminished without a statutory clause in which Congress explicitly declares its intent, but whether such a presumption even exists.

One author points to the apparent absurdity of an instance in which an entire omnibus spending statute could have been invalidated for the sake of a single invalid provision.[380] Such an instance is a textbook example of where severability is appropriate. The Constitution requires that money from the federal treasury can only be spent as a result of Congress enacting an appropriations bill.[381] Congress normally funds the government through thirteen separate appropriations bills every year, each covering certain departments and agencies.[382] An omnibus appropriations bill is one that combines what would normally be multiple annual spending bills into a single bill.[383] Appropriations bills typically include "riders," which are instructions on the spending, such as a prohibition on the

---

could allow for the possibility that the legislative act of removing the clause could still be considered weak evidence of intent, the court of appeals seems to be saying that it was not evidential to any extent. The circuit court went on to make the shocking statement that only an express nonseverability clause in the final version of the bill should impact a severability analysis. *See id.* at 1322–23. Although this aspect of the court's holding is manifestly wrong under current severability doctrine, as demonstrated throughout this Article, its flawed analysis aside, the fact remains that removal of a severability clause does not factor into severability analyses in that circuit.

Requiring an express nonseverability clause is unrealistic as a political matter, so if courts require such a clause to hold a statute nonseverable then very few statutes will be nonseverable. Opponents of controversial legislation might occasionally attempt to insert such a clause, but it is unlikely such attempts will ever succeed, since if there are sufficient votes to insert the clause, there should often be sufficient votes to defeat the legislation. *See, e.g.*, Schumsky, *supra* note 183, at 229–30 (citing, inter alia, 147 CONG. REC. S3084, S3088–90 (daily ed. Mar. 29, 2001) (statement of Sen. Frist)). Although there are federal statutes with nonseverability clauses, *e.g.*, Mobile Telecommunications Sourcing Act, Pub. L. No. 106-252 § 125, 114 Stat. 626, 632 (2000) (codified at 4 U.S.C. § 125 (2006)); Catawba Indian Tribe of South Carolina Land Claims Settlement Act of 1993, Pub. L. No. 103-116 § 15, 107 Stat. 1118, 1136 (1993) (codified at 25 U.S.C. § 941(m) (2006)), such clauses are rare. *See* Schumsky, *supra* note 183, at 243–44.

380. *See* Nagle, *supra* note 5, at 204 & n.1.

381. U.S. CONST. art. I, § 9, cl. 7.

382. *See generally* Sandy Streeter, The Congressional Appropriations Process: An Introduction, Congressional Research Service (Dec. 2, 2008), *available at* http://assets.opencrs.com/rpts/97-684_20081202.pdf (indicating that the House and Senate usually consider twelve regular appropriations bills and one omnibus appropriations bill).

383. *Id.* at 5–10. This usually occurs when Congress fails to fulfill its constitutional duty to pass appropriations bills in a timely manner, and as a result pass an all-encompassing measure to provide all remaining spending for the fiscal year.

appropriated funds being spent for a particular purpose.[384] In 1988, the D.C. Circuit considered the constitutionality of such a rider provision in an omnibus appropriations bill.[385] The court invalidated one provision in the rider (not even the whole rider),[386] but in doing so severed it from the remainder of the statute,[387] upholding the vast majority of the legislation.

This is an instance where one could argue the proper judicial remedy is implied severability. The omnibus legislation (styled as a continuing resolution) was a 471-page statute.[388] Appropriations bills are to fund all the operations and programs of the federal government. The D.C. Circuit ruled that:

> The question of whether one part of a statute is severable from another is primarily one of legislative intent, informed by a general presumption of severability. Although the two parts of the [challenged appropriations rider] are tangentially related, we see no indication that Congress would not have enacted the first part of the amendment without the second.[389]

Despite occasional brinksmanship, rarely would Congress rather shut down the entire United States government than pass an omnibus sans one-half of one rider in a lengthy bill.

Thus, the absence of a severability clause lowers the burden that must be met to hold the entire statute severable, and the question becomes the proper conceptualization of these standards. The *Alaska Airlines* Court held that a severability clause means a provision is severable unless there is "strong evidence" that Congress intended otherwise.[390] By negative inference, this suggests that the evidence need not be "strong" in the absence of a severability clause. The challenge becomes

384. *Cf.* BLACK'S LAW DICTIONARY 1342 (8th ed. 2004) (defining "rider" as "[a]n attachment to some document, such as a legislative bill or an insurance policy, that amends or supplements the document. A rider to a legislative bill often addresses subject matter unrelated to the main purpose of the bill").

385. News Am. Pub., Inc. v. FCC, 844 F.2d 800, 801 (D.C. Cir. 1988).

386. *See id.* at 815–16.

387. *Id.* at 802 n.1, 815.

388. *Id.* at 801.

389. *Id.* at 802 n.1 (citing Regan v. Time, Inc., 468 U.S. 641, 653 (1984); Buckley v. Valeo, 424 U.S. 1, 108 (1976)). This citation to *Buckley* is puzzling, however. There is nothing in the cited portion of *Buckley* supporting the D.C. Circuit's holding on this point, nor does Judge Stephen Williams quote any relevant language in this citation.

390. 480 U.S. 678, 686 (1987). This invites cynical comments about Congress not meaning what it said.

how to codify that suggestion in a standard of proof that courts can reliably apply.

The Court's language in *Alaska Airlines* raises two possibilities. The first is that the law is still presumed severable, but with the presumption weakened. As previously discussed, this is the position adopted by several circuits but never by the Supreme Court. Under this theory, less persuasive evidence is required to establish that the provision is nonseverable. This would be analogous to different burdens of proof for triers of fact. Proof by a preponderance of the evidence merely means that one possibility is more likely than the other. Otherwise put, it means that the probability of a particular outcome in a binary inquiry is greater than 50%, rendering the probability of the opposite outcome less than 50%.

The second is that there is no presumption either for or against severability. This is the better position. This would mean that the judicial predisposition of two opposing possibilities to a binary question stand in precise equipoise. Simply put, the court presumes nothing.[391] The burden remains on the challenger to argue nonseverability merely by virtue of the fact that there is no presumption against the government, and so as a matter of inertia the *status quo ante* (of an enacted statute being implemented) continues unless the challenger provides the court with an argument that alters the situation. The burden of persuasion is different from the burden of proof, in that the former asks who must make the argument, and the latter asks how strong of a showing the arguer must make.[392] Therefore

391. This appears to be the position suggested by the D.C. Circuit in an opinion by Judge Douglas Ginsburg, *see* Bismullah v. Gates, 551 F.3d 1068, 1073 (D.C. Cir. 2009) ("Congress's failure to include a non-severability clause does not create a presumption of severability, any more than the absence of a severability clause implies non-severability."), despite precedent to the contrary, City of New Haven v. United States, 809 F.2d 900, 905 (D.C. Cir. 1987).

392. As these two burdens are often confused, it bears noting that there are issues wherein the party that bears the burden of proof need not show that the evidence supports the bearer's position, whereby the bearer can prevail even when the evidence weighs against the bearer. For example, several areas of law require a showing of substantial evidence—which is considerably less than a preponderance of the evidence—as sufficient for certain issues. In administrative law, on questions of fact or policy challenging agency action, substantial evidence is the standard required under the Administrative Procedure Act (APA), *see* Pub. L. No. 79-404 § 10(e), 60 Stat. 237, 243 (1946) (codified at 5 U.S.C. § 706(2)(E) (2006)). It does not require the weight of the evidence to support the matter asserted. "Substantial evidence is more than a scintilla, and must do more than create a suspicion of the existence of the fact to be established." NLRB v. Columbian Enameling & Stamping Co., 306 U.S. 292, 300 (1939), *cited in* Allentown Mack Sales & Serv. v. NLRB, 522 U.S. 359, 367 (1998). *See also* FED. R. EVID.

without the burden shifting onto the defendant, it remains where it originally was.

It is possible that the most reasonable standard of proof under *Free Enterprise* Step Two when a severability clause is missing is a preponderance of the evidence. The standard must be less rigorous than clear and convincing evidence, since as shown above clear and convincing evidence is the requirement when a severability clause is present, and the Supreme Court has made clear that the burden is lessened without the clause. It is improbable—though still possible—that the standard cannot merely be substantial evidence or any other threshold less than fifty percent probability, because the Court had more recently indicated it must be "likely" that Congress would not want the abridged statute to remain in force, not merely plausible or possible that Congress might accept the shortened statute.[393]

A preponderance standard might strike this balance, though earlier Supreme Court cases antedating *Alaska Airlines* (and also *Champlin*, for that matter) suggest something akin to a "substantial evidence" standard. Beyond being less demanding than a preponderance of the evidence, this earlier standard merely required significant doubt about whether Congress would have passed the truncated statute.[394] It would also ironically assign the judiciary a more modest role if the standard was something akin to substantial evidence, as courts would send statutes back to Congress whenever there is significant evidence that Congress might not be pleased with the remaining statute when Congress does not declare any intent in the statute's text. The proper standard thus seems a bit uncertain, so Supreme Court elaboration on this point would be quite helpful.

It is possible that a court need not decide which of two possibilities (weakened presumption versus no presumption)

---

104(b); *cf.* Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 248, 252 (1986) (defining a genuine issue of material fact sufficient to defeat a motion for summary judgment as one in which there is more than a scintilla of evidence favoring the nonmoving party, such that a jury could reasonably find for the nonmoving party). So the evidence must be more than insubstantial, but nevertheless can be far less than a preponderance.

393. *See* United States v. Booker, 543 U.S. 220, 249 (2005) (requiring courts to inquire if "the scheme that Congress created" would be "so transform[ed]" without the invalid provision "that Congress *likely* would not have intended the Act as so modified to stand") (emphasis added).

394. *See* El Paso & Ne. Ry. Co. v. Gutierrez, 215 U.S. 87, 97 (1909) ("If we are satisfied that it would not, or that the matter is in such doubt that we are unable to say what Congress would have done omitting the unconstitutional feature, then the statute must fall.").

accurately describes a situation involving a statute without a severability clause, because there is little practical difference between conceptualizing the inquiry as the former, rather than the latter. If the challenger must carry the burden of proof, and that burden is not satisfied by carrying the requisite burden of proof (whether a preponderance of the evidence, or substantial evidence) then, regardless of what the presumption was, the *status quo ante* endures, and so only the invalid portion would be stricken from the statute. Nevertheless, the Supreme Court should clarify this matter to provide guidance to lower courts and to allow lawyers and scholars to correctly frame legal arguments.

Presuming a statute to be severable and tasking the challenger with rebutting that assumption shows due respect for the democratic lawmaking process. However, courts must be careful not to set too high a bar for challengers to meet, as this would directly run afoul of the Supreme Court's contrary indications that legislatures cannot simply cast a large net of legislation containing constitutionally dubious provisions, and task the judiciary with separating the wheat from the chaff.[395] Then again, presuming a statute nonseverable can also be cited as judicial modesty, since a court removes itself from the business of second-guessing how much of a statute Congress would be satisfied to retain. This caveat of not giving legislatures too much leeway is also reinforced by the fact that many state courts still apply a presumption of nonseverability to state statutes lacking a severability clause.[396]

### B. *Extratextual Severability Clauses*

It is also an unsettled question as to whether a severability clause can only be effective if contained within the four corners of a statute as enacted, meaning in the precise form that it was enrolled in Congress, presented to the President, and signed by the President as required by the Constitution (or re-passed by a two-thirds congressional vote to override a presidential veto).[397] If a court holds part of a statute invalid, and that statute lacks a

395. *See* Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 330 (2006) (quoting United States v. Reese, 92 U.S. 214, 221 (1876)).

396. Nagle, *supra* note 5, at 220 n.86 (citing as an example State v. Aldrich, 231 N.W.2d 890, 895 (Iowa 1975)).

397. U.S. CONST. art. I, § 7, cl. 2.

severability clause, but the statute amends a previous statute which did contain a severability clause, can the court credit this earlier clause to also cover the later amending statute? The Court in *Alaska Airlines* expressly disclaimed this question, as the Court was examining a challenge to a statute lacking a severability clause, but which in turn amended a prior statute that did contain such a clause.[398] The Court seemed skeptical of this approach, but nonetheless did not foreclose the possibility.[399] A corollary theory would be whether a subsequent enactment containing a severability clause could retroactively insert severability into a prior statute, where the later statute expressly announces that it amends the previous Act.

Such a possibility is both counterintuitive and exceedingly ill-advised under separation-of-powers principles. The enactment of a statute is a singular event, wherein both Houses of Congress pass a measure and the President signs it. The moment of the House vote is the moment wherein the will of the American people is expressed. The moment of the Senate vote is the moment wherein the will of the states is expressed. And the moment of the President's signature is the moment wherein the will of the head of state is expressed. There is no way to perfectly reconstruct the precise combination of thoughts and intentions of 435 Representatives, 100 Senators, and one President in determining whether the overall legislative bargain embodied in the bill before them was deserving of their acceptance or their rejection. Severability does not apply to the *United States Code*, into which all of these statutes flow; it is found in the *Statutes at Large*, and rather applies to each discrete enactment which is the tangible result of the democratic process. If the legislature at the moment of enactment expresses a wish that the provisions of the particular bill currently before it are severable, then so be it; but if not, then the legislature declined to make such a declaration, and instead decided to pass the measure without expressing any such wish. As the field of ascertaining legislative intent is already fraught with uncertainty, it injects an unacceptable level of potential revisionism to empower courts to depart from the text

398. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 686–87 & n.8 (1987) (noting that the Airline Deregulation Act of 1978, then at bar, amended the Federal Aviation Act of 1958, which contained a severability clause. *See* Pub. L. No. 85-726 § 1504, 72 Stat. 731, 811 (1958) (codified at 49 U.S.C. § 1301 note)).

399. *Id.* at 687 n.8.

of an enactment, and import language from a subsequent enactment which—although concerning the same subject matter—is nonetheless not a recreation of the legislative will at the moment of the adoption of the original statute.

### C. *Congruous with Other Areas of Law*

Two other areas of law have direct application to severability. Together, these two areas show a methodological consistency which reflects a common conceptual approach in statutory interpretation.

#### 1. *Chevron* Deference Versus *Skidmore* Deference

Setting the bar in severability inquiries in two different places—one for statutes with a severability clause and one for those without—is analogous to judges' approach in other areas of law. Administrative law contains an excellent example of employing two different rules in statutory interpretation. Under *Chevron* deference, when an administrative agency has fulfilled a specific rigorous fact-finding and policy-making process, courts defer to the agency's interpretation of ambiguities or filling in a gap on interstitial issues where the statute is silent so long as the agency's construction is reasonable.[400]

But the Supreme Court held in 2001 that *Chevron* deference does not attend all agency interpretations of statutes.[401] Instead, the Court held *Chevron* only applies if Congress has "delegated authority to the agency generally to make rules carrying the force of law, and . . . the agency interpretation claiming deference was promulgated in the exercise of that authority."[402] Thus *Chevron* deference attaches only if certain criteria are met, such as when the agency's interpretation is the result of notice-and-comment rulemaking or formal administrative adjudication.[403] The Court has enumerated various agency pronouncements that fall short of this standard.[404]

400. Chevron U.S.A. Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837, 842–43 (1984). An agency can also change its policy position so long as the agency presents a reasoned analysis. Nat'l Cable & Telecomms. Ass'n v. FCC, 567 F.3d 659, 667 (D.C. Cir. 2009) (quoting Motor Vehicle Mfrs. Ass'n v. State Farm, 463 U.S. 29, 57 (1983)).

401. *See* United States v. Mead Corp., 533 U.S. 218, 227–31 (2001).

402. *Id.* at 226–27.

403. *Id.* at 227; *accord* Lisa Schultz Bressman, *How* Mead *Has Muddled Judicial Review of Agency Action*, 58 VAND. L. REV. 1443, 1444–45 (2005). The Court has not yet issued an exhaustive list of when *Chevron* applies. It is not necessarily limited to formal rulemaking,

Even when *Chevron* deference does not apply, however, agency interpretations are still given an inferior level of deference.[405] Agency interpretations that do not satisfy requirements analogous to the rigorous standards involved in notice-and-comment rulemaking or quasi-judicial adjudications are still afforded *Skidmore* deference,[406] under which interpretations are adopted if they are persuasive.[407]

There are three parallels to severability doctrine as portrayed in this Article. The first is that *Chevron* is a two-step framework, explicitly articulated and applied in that fashion.[408] With Step One applying when statutes clearly speak to the question at issue and a court proceeding to Step Two only when the statute is silent or ambiguous. This was actually my inspiration for characterizing *Free Enterprise* as creating a two-step analysis for severability, under which Step One examines whether the statute's remaining provisions are still operable, and, only if they

---

Pub. Citizen, Inc. v. HHS, 323 F.3d 654, 660 (D.C. Cir. 2003) (citing Barnhart v. Walton, 535 U.S. 212, 221 (2002); *Mead*, 533 U.S. at 231)), and at least three scholars argue that *Chevron* should be limited to the two examples thus far authorized (i.e., formal rulemaking and adjudications), *see* Thomas W. Merrill & Kristin E. Hickman, Chevron's *Domain*, 89 GEO. L.J. 833, 884–85 (2001); Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461, 541–44 (2003).

404. *Mead*, 533 U.S. at 234 (policy statements, agency manuals, and enforcement guidelines); Christensen v. Harris Cnty., 529 U.S. 576, 587 (2000) (same); *see also* Wash. State Dep't of Soc. & Health Servs. v. Guardianship Estate of Keffeler, 537 U.S. 371 (2003) (holding that *Chevron* does not apply to Social Security Act interpretations contained in that agency's operational manual).

405. *Chevron* deference is not to be confused with *Auer* deference. An agency is given significance deference when it issues definitions of its own regulations (that is, issues arising from legally-binding interpretations of regulations formerly promulgated by that same agency). Such a regulatory interpretation is "controlling unless plainly erroneous or inconsistent with the regulation" or if there is any reason to doubt that the agency's views reflect a fair and considered judgment. Auer v. Robbins, 519 U.S. 452, 461, 462 (1997) (internal quotation marks omitted). The former (*Chevron* deference) concerns an agency's interpretation of statutes. The latter (*Auer* deference) concerns an agency's interpretation of regulations.

406. *See* Skidmore v. Swift & Co., 323 U.S. 134, 140 (1944) (noting that agency decisions that do not have the power to control may still be persuasive).

407. *Mead*, 533 U.S. at 221 (holding that *Skidmore* deference renders an agency interpretation "eligible to claim respect according to its persuasiveness); *see also Christensen*, 529 U.S. at 587 (holding that when *Chevron* deference is unjustified statutory interpretations are still "entitled to respect . . . but only to the extent that those interpretations have the power to persuade" (quoting *Skidmore*, 323 U.S. at 140) (internal quotation marks omitted)).

408. *See, e.g.*, Sherley v. Sebelius, 644 F.3d 388, 399 (D.C. Cir. 2011) (reversing the district court's grant of a preliminary injunction on a stem-cell funding program); Anna Jaques Hosp. v. Sebelius, 583 F.3d 1, 2 (D.C. Cir. 2009) (sustaining an agency interpretation of the Medicare reimbursement statute under *Chevron* Step Two); *Pub. Citizen*, 323 F.3d at 658–59 (declining to apply either *Chevron* Step One or Step Two to the agency interpretation at issue).

are, does a court proceed to Step Two to consider whether the statute still functions in the manner Congress intended.[409]

The second parallel is the role of textual interpretation. The first step of each is exclusively based on the text of the statute at issue. Under *Free Enterprise*, the question is whether the statutory provisions can still functionally interact according to their text. Under *Chevron*, the question is whether the statute's text is unambiguous on the issue at bar. If this textual approach is insufficient, the court must then move on to a step that is less objective and relies more heavily on judicial reasoning. Under *Free Enterprise*, Step Two asks whether Congress's purposes are still fulfilled by the statute functioning in the manner Congress intended. Under *Chevron*, Step Two is whether the agency interpretation is reasonable.

The third parallel is that both determine which of two analytical frameworks applies based on one threshold question. For agency interpretations, whether *Chevron* deference applies versus *Skidmore* deference depends on one criterion (i.e., whether certain procedural formalities have been satisfied). For severability, the criterion is whether the words of the statute are still coherent without the invalid provision. So for both, a court begins with one standard setting a particular bar for court action, and based on whether one threshold criterion is satisfied a court may then conduct its analysis on a separate track. And in both instances, the second track sets a less rigorous standard that must be satisfied before the court can trump the decision of a coequal branch (overcoming Executive Branch interpretations under *Chevron* and *Skidmore*, and overcoming Congress's statutory enactments under *Free Enterprise*).

This deferential framework approach has worked reasonably well for administrative law, criticisms and contrary predictions notwithstanding.[410] It also reflects an understanding that while a single rule works for many areas of law, a two-tier approach works better in other areas. This is true for courts evaluating

409. Frameworks wherein clear statutory text trumps all extrinsic factors are also consistent with the textualist and originalist schools of thoughts examined in Part II.C, *supra*. *See also* Eskridge, *New Textualism*, *supra* note 236, at 622–23.

410. Justice Antonin Scalia was the sole dissent in the Supreme Court's 8–1 decision in *Mead*, in which he predicted that introducing an alternative standard to the two-step *Chevron* inquiry would create disastrous problems for the judiciary. *See Mead*, 533 U.S. at 241 (Scalia, J., dissenting). Some scholars have subsequently echoed those concerns to one extent or another. *See, e.g.*, Bressman, *Agency Action*, *supra* note 403, at 1444.

agency interpretations of the statutes they implement, and it is true for courts determining how much of a statute to throw out with an unconstitutional provision.

2. Preemption Doctrine Considers Overall Legislative Scheme

The other doctrine characterized by congruencies with severability is preemption doctrine, arising from the Supremacy Clause of the Constitution,[411] under which federal law trumps conflicting state laws.[412] Preemption issues exist between the federal government and a state when the laws of the former clash with those of the latter. "Where state and federal law directly conflict, state law must give way."[413] In such situations, courts try to take care in navigating confrontations between these dual sovereigns by finding preemption when Congress clearly expresses its intent to override state law, especially when the subject matter at issue is one traditionally managed by the states.[414] In one form of conflict preemption—called obstacle preemption—such a conflict exists between the two forms of government when a state law "stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress."[415] Another form of conflict

411. U.S. CONST. art. VI, cl. 2.

412. Crosby v. Nat'l Foreign Trade Council, 530 U.S. 363, 372 (2000).

413. PLIVA, Inc. v. Mensing, 131 S. Ct. 2567, 2577 (2011) (internal quotation marks omitted). The Court has also made clear what constitutes federal and state law directly conflicting. The Court has held "that state and federal law conflict where it is impossible for a private party to comply with both state and federal requirements." *Id.* (quoting Freightliner Corp. v. Myrick, 514 U.S. 280, 287 (1995)) (internal quotation marks omitted).

414. Federal courts:

> must be guided by two cornerstones of [Supreme Court] pre-emption jurisprudence. First, the purpose of Congress is the ultimate touchstone in every pre-emption case. Second, [i]n all pre-emption cases, and particularly in those in which Congress has legislated . . . in a field which the States have traditionally occupied, . . . [federal courts] start with the assumption that the historic police powers of the States were not to be superseded by the Federal Act unless that was the clear and manifest purpose of Congress.

Wyeth v. Levine, 555 U.S. 555, 565 (2009) (quoting Medtronic, Inc. v. Lohr, 518 U.S. 470, 485 (1996)) (internal quotation marks and citations omitted).

This preemption can take one of two forms, field preemption and conflict preemption. "When Congress intends federal law to occupy" a particular subject matter, "state law in that area is preempted." *Crosby*, 530 U.S. at 372 (quoting California v. ARC Am. Corp., 490 U.S. 93, 100 (1989)) (internal quotation marks omitted). When Congress does not intend to block state legislation so broadly as to permeate the subject matter, "even if Congress has not occupied the field, state law is naturally preempted to the extent of any conflict with a federal statute." *Id.*

415. *Crosby*, 530 at 373 (quoting Hines v. Davidowitz, 312 U.S. 52, 67 (1941)).

preemption—actual conflict—is when federal and state laws directly conflict, and exists for citizens when, as a result of contradictory laws, it is "impossible for a private party to comply with both state and federal requirements."[416]

As with the administrative-law issue explored above, there are congruencies between severability doctrine and preemption doctrine. The first is that, just as with severability, the touchstone in preemption doctrine is ascertaining congressional intent.[417] Just as the Supreme Court has expounded on various methods for assessing congressional intent discussed in this Article, the Court has likewise prescribed the same approach for preemption. And just as the single greatest indicium of severability intent is an express severability clause, so too the strongest indicia of congressional preemptive intent are those found in the statute's text.[418]

The second parallel is closely related: Courts must assess Congress's intent in the statute as an integrated whole and determine the implications of the contemplated judicial action vis-à-vis the overall statutory scheme embodied in the enactment. This is central to *Free Enterprise* Step Two and was the core holding of *Alaska Airlines*.[419] Likewise in one modern preemption case, the Supreme Court found that the challenged statute "exert[s] an extraneous pull on the scheme established by Congress" to alter a carefully-balanced legislative bargain.[420] For each statutory provision, courts must "focus on the plain wording of the clause, which necessarily contains the best evidence of Congress' preemptive intent."[421] The Court's language in these cases mirrors *Alaska Airlines*, where the Court held that judges must determine the "original legislative bargain" Congress struck in formulating a statute, and can only sever an invalid provision to retain the remaining statute if the general governing dynamics of the statute can still function in a manner consistent with Congress's intent.[422]

416. Freightliner Corp. v. Myrick, 514 U.S. 280, 287 (1995) (internal quotation marks omitted).

417. *See* cases cited *supra* note 414.

418. *See, e.g.*, *Wyeth*, 555 U.S. at 566–67.

419. *See* Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987); *also supra* Parts II.B.1.b & IV.A.2.

420. Buckman Co. v. Plaintiffs' Legal Comm., 531 U.S. 341, 353 (2001).

421. CSX Transp., Inc. v. Easterwood, 507 U.S. 658, 664 (1993).

422. *Alaska Airlines*, 480 U.S. at 685.

Just as with administrative deference, courts should likewise be able to import interpretive methods from preemption case law into severability. It is a preemption "principle that it is Congress rather than the courts that pre-empts state law."[423] Likewise, it is Congress's intentions, not the preferences of the courts one way or the other, that controls whether an invalid provision can be severed from a statute without disrupting the legislative bargain that catalyzed the statute's creation. One of the greatest challenges in severability cases is correctly conceptualizing Congress's overall statutory scheme to determine whether the statute can still function in the manner Congress intended without the invalid provision. Being able to analogize to and distinguish from preemption cases engaging in the same inquiry would both help develop objective methods that produce reliable outcomes for this elusive aspect of severability inquiries, and also further harmonize severability doctrine with other interpretive doctrines.

## VI. SEVERABILITY AS A DOCTRINE OF JUDICIAL RESTRAINT

Severability is fundamentally a doctrine of judicial restraint. In accordance with the principles and rules already explored in this Article, it proceeds from a judicial modesty that seeks to retain as much of a statute as practicable, stemming from a recognition that, to the extent that an enactment is invalidated, a court undoes the tangible product of the democratic process. "Judicial restraint counsels against striking down an entire piece of legislation on the basis of some constitutional infirmity in a minor provision."[424] As unelected officials in a system of government wherein officials are generally accountable to the electorate, judges rightly prefer to strike down unconstitutional enactments only insofar as necessary to reconcile such enactments with the Constitution.

Judicial restraint does not always mean severing an unconstitutional provision from the remainder of the statute. As shown below, sometimes proper restraint requires striking down a substantial part of a statute. And counterintuitive though it

423. Gade v. Nat'l Solid Wastes Mgmt. Ass'n, 505 U.S. 88, 111 (1992) (Kennedy, J., concurring in part and concurring in judgment).

424. Noah, *supra* note 26, at 236–37. By specifying this sentiment should apply to *minor* provisions, the author (by negative inference) suggests the contrapositive is true—that judicial restraint does not discourage severability when the invalid provision is a major provision in the statute.

may seem, still other times judicial focus on restraint and modesty requires a court to invalidate a statute in its entirety.

### A. *Severability is Premised on Judicial Restraint*

The doctrine governing severability is a self-imposed concept to constrain judicial power. Discussing severability, the Supreme Court recently began by stating, "Generally speaking, when confronting a constitutional flaw in a statute, we try to limit the solution to the problem."[425] This expresses prudent reticence on the part of judges, and of courts recognizing both that their institutional competence is limited in policymaking, and also that propriety requires their governmental role be limited in a democratic society.[426]

When encountering an unconstitutional aspect of a statute in a constitutional challenge, courts must make a surgical determination. Like a surgeon discovering cancer in a patient on the operating table, the physician must assess how far the cancer reaches, taking care to remove the malignant tissue while saving as much healthy tissue as possible. This surgical metaphor aptly describes the judge's task, in that the judicial mandate is to respect and retain, to the extent possible, the product of the democratic process.[427]

Two longstanding judicial rules also sound in severability doctrine, first that courts are "never to anticipate a question of constitutional law in advance of the necessity of deciding it; the other never to formulate a rule of constitutional law broader than is required by the precise facts to which it is to be applied."[428] These rules reinforce the concept of judges using the power of judicial review sparingly and circumspectly. They also suggest that severability doctrine can be better understood by consulting the rationales underlying other doctrines of

425. Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 328 (2006).

426. This is analogous to the view that courts should be mindful of America's free-market economic system when interpreting regulatory statutes, and not assign to those statutes an unnecessarily broad sweep that unduly undermines the ability of private parties to form contracts and conduct business matters. *See* JEREMY RABKIN, JUDICIAL COMPULSIONS: HOW PUBLIC LAW DISTORTS PUBLIC POLICY (1989); Frank H. Easterbrook, *Statute's Domains*, 50 U. CHI. L. REV. 533, 544–51 (1983).

427. *See supra* Parts III.B (discussing *Ayotte*, 546 U.S. at 328–30).

428. United States v. Raines, 362 U.S. 17, 21 (1960). Although *Raines* was not a severability case, and thus the Court was reiterating these rules in another context, the Court also invoked these rules in a subsequent case that was considering severability. *See* Brockett v. Spokane Arcades, Inc., 472 U.S. 491, 501 (1985).

judicial restraint, harmonizing them to the extent possible so as to demonstrate that they arise from a common conceptual framework.

It is sometimes false humility when judges find an invalid provision severable. The reality is that sometimes judicial modesty requires totally severing invalid provisions. Other times judicial restraint requires partial severability.[429] Still, other times the most modest approach is to hold the statute completely nonseverable. Courts must be mindful of these various possibilities and reject a kneejerk tendency to strike down only the inherently flawed provision. Consider how those three outcomes—total severability, partial severability, and total nonseverability—can result from the three principles from *Ayotte* discussed in Part III when applied to different types of statutes.

This holds true regardless of whether a court is applying *Free Enterprise* Step One or Step Two.[430] This inquiry is not an overly-taxing exercise when a court is in *Free Enterprise* Step One, as the court must simply consider whether any parts of the statute malfunction or dysfunction without the excised provision. If the invalid provision is literally essential to part of the statute, then only that part must fall along with the invalid provision, while if the invalid provision is functionally vital to the entire statutory scheme then the entire statute must fall.

A court's work becomes considerably more complex if a court moves on to *Free Enterprise* Step Two. A court must ascertain the "original legislative bargain" embodied by the statute,[431] and then determine whether the statute can still function "in a *manner* consistent with the intent of Congress."[432] But on many occasions, a statute can be subdivided into different legislative bargains, similar (though not identical) to how a statute can be separated into self-contained functional units considered under Step One. The obvious difference is that Congress ultimately votes on a statute as one document in a take-it-or-leave-it proposition, complicating the task of determining whether Congress would have passed the part containing the invalid provision had it known the provision would later be stricken. Often indications of congressional intent come from statements

429. Partial severability was explained in Part IV.B.1, *supra.*

430. *See supra* Part IV.A.

431. Alaska Airlines, Inc., v. Brock, 480 U.S. 678, 685 (1987).

432. *Id.*

or actions pertaining to the entire statute, and such indications rarely draw the line of particular subdivisions of the statute when illuminating Congress's purposes.

These challenges notwithstanding, partial severability is sometimes the more restrained approach for a court to implement than either total severability or total nonseverability. Take, hypothetically, a statute that is both large and complex. It is organized into four chapters. One section of the statute that is found in Chapter 2 is held constitutionally infirm. That section is quite clearly indispensible to Chapter 2 in which it is located, as the central component in an interrelated scheme. But each of the four chapters is reasonably independent of the others. In such a situation, the correct result would likely be to hold the invalid section partially severable, requiring the invalidation of Chapter 2 but allowing Chapters 1, 3, and 4 to survive the challenge.

### B. *Justices Rehnquist and White: Example of Judicial Restraint Requiring Complete Invalidation*

One relatively recent case is quite revealing on where the proper line should be drawn for severability among those who advocate for judicial restraint. The most recent example of Justices voting to strike down an entire statute as nonseverable is Justice William Rehnquist's dissent in *INS v. Chadha*.[433]

This dissent reveals the fallaciousness of the argument that judges who emphasize judicial restraint must vote for severability, especially when a severability clause is present. Justice Rehnquist was the most conservative member of the Supreme Court in 1983, and it is fair to characterize Justice Byron White, who joined Justice Rehnquist's dissent, as one of the more conservative members of the Court at that time, though they came from different legal generations, with different judicial philosophies.

The *Chadha* case was not particularly controversial. It merely involved the administrative procedures and governmental options associated with the deportation of noncitizens.[434] Also, the implications for the case were not widespread, as it only concerned legislative-veto provisions in a number of federal

433. 462 U.S. 919, 1013–16 (1983) (Rehnquist, J., joined by White, J., dissenting).

434. *See id.* at 923–24 (majority opinion).

statutes.[435] So there can be no credible accusation that any of the opinions in *Chadha* were driven by an *ad hoc* approach to reach some preordained result.

It is not surprising that Rehnquist's dissent in *Chadha* has received relatively little attention, his stature as one of the more intelligent and consequential members in the history of the Court notwithstanding. Whereas *Chadha* was a 6–3 decision, *Alaska Airlines* was a unanimous decision of the Court decided only four short years later, and elaborated on severability doctrine in much more detail. Thus, *Alaska Airlines* predictably eclipsed *Chadha*, and so much judicial and scholarly focus was directed at *Alaska Airlines* that it is unsurprising that a dissent from *Chadha* has been largely overlooked.

Justice William Rehnquist argued that the legislative veto was nonseverable from the remainder of the Immigration and Nationality Act of 1952, and thus that the entire statute should be struck down.[436] Rehnquist wrote that Congress likely did not intend to allow the Attorney General to suspend deportations if Congress lacked the power to override his decision.[437] Quoting from an earlier case, Rehnquist explained that severing the invalid provision creates the result that, "the statute is made to enact what confessedly the Legislature never meant. It confers upon the statute a positive operation beyond the legislative intent, and beyond what anyone can say it would have enacted in view of the illegality of [the challenged provision]."[438]

Rehnquist noted that the legislative-veto provision was an excepting clause, and went on to specifically add that the Court's reasoning concerning excepting clauses reinforced his conclusion.[439] As Rehnquist quoted from yet another case, "Where an excepting provision in a statute is found unconstitutional, courts very generally hold that this does not work an enlargement of the scope or operation of other provisions with which that provision was enacted and which was intended to qualify or restrain."[440]

435. *But see id.* at 959–60 (Powell, J., concurring) (arguing the number of statutes being modified by the Court's holding was significant, and thus the Court's holding was broad and consequential); *id.* at 967 (White, J., dissenting) (same).

436. *Id.* at 1014 (Rehnquist, J., dissenting).

437. *Id.* at 1014.

438. *Id.* (quoting Spraigue v. Thompson, 118 U.S. 90, 95 (1886)).

439. *Chadha*, 462 U.S. at 1014–15 (Rehnquist, J., dissenting).

440. *Id.* at 1014 (quoting Davis v. Wallace, 257 U.S. 478, 484 (1922)).

But Rehnquist (and with him Byron White) then made a much broader statement that should apply in all severability cases. The future Chief Justice argued that courts must recognize legislative intent as derivable from all the provisions of the statute, and consider whether this intent is still being served by the abridged statute.[441] Rehnquist continued quoting from this case, which in turn adopted the rule from an Ohio severability case.[442] The Supreme Court incorporated the Ohio court's reasoning, writing of refusing to honor the intent of the invalid provision:

> This would . . . mutilate the section, and garble its meaning. The legislative intention must not be confounded with their power to carry that intention into effect. To refuse to give force and vitality to a provision of law is one thing, and to refuse to read it is a very different thing. It is by a mere figure of speech that we say an unconstitutional provision is 'stricken out.' For all the purposes of construction it is to be regarded as part of the act. The meaning of the legislature must be gathered from all they have said, as well from that which is ineffective for want of power, as from that which is authorized by law.[443]

Noting that the majority found Congress wanted the legislative-veto provision to be severable because Congress wanted to lessen its workload, Rehnquist and White rejected the majority's assessment, arguing to the contrary that legislative history showed that Congress wanted to deny the Executive Branch unilateral control of deportation suspensions.[444] Rehnquist went on to note that there were other legislative formulations by which Congress could advance the same goal (of restraining the Executive's suspension authority), but that Congress's declination to do so meant the Court should not presume to conclude Congress would prefer the remaining statute to continue in force.[445] This last point is quite interesting in that Rehnquist specifically added that it is not a federal judge's role to make such a determination unless there are affirmative

441. *Id.* at 1015.

442. *Id.* at 1014–15.

443. State *ex rel.* McNeal v. Dombaugh, 20 Ohio St. 167, 174 (1870), *quoted in Davis*, 257 U.S. at 484–85.

444. *Chadha*, 462 U.S. at 1015 (Rehnquist, J., dissenting).

445. *Id.* at 1016.

indications of congressional intent to advance the statute's policy objectives without the invalid provision.[446]

In other words, Rehnquist and White specifically argued that proper judicial restraint requires not only that a court inquire as to whether the statute's overall objective would be advanced without the invalid provision. Instead, a court must also not indulge in the presumption that Congress would be satisfied with a rebalanced statutory scheme absent some evidence of Congress's intent in that narrow regard. Unless the record contains evidence suggesting Congress would be willing to allow a reformulated operative scheme, a court should hold the provision nonseverable even if the remaining statute still moves in the direction of advancing the statutory purpose. Thus, Rehnquist concluded, by severing the legislative-veto provision and retaining the rest of the statute, the majority "ha[d] confounded Congress's intention . . . with their power to carry that intention into effect."[447] Indeed, at least two federal appeals courts evidently agree with Rehnquist as to where the line should be drawn, as they held two statutes with legislative-veto provisions nonseverable, invalidating them *in toto*.[448]

What makes this dissent particularly interesting is that—as already noted in this Article—the Immigration and Nationality Act at issue in *Chadha* contained a severability clause. Thus even with an express indication of congressional intent, Rehnquist and White both believed the legislative veto in the statute to be so significant to the statute as a whole that Congress would not have wanted the remaining statute to continue in effect without it—that the policy formulated by the statute should instead be returned to Congress to recalibrate the statute's provisions in light of the fact that Congress could not retain a veto-like power over the Attorney General on deportations. Such considerations become even more important after the Court decided *Alaska Airlines* four years later with the focus on a statute functioning in the manner Congress intended, and suggests the judicial thinking Rehnquist and White carried with them into *Alaska*

446. *See id.* 1015–16 (criticizing the majority for severing the Act despite having no indication that Congress wished for it to take effect in its severed form).

447. *Id.* at 1016 (quoting *Davis*, 257 U.S. at 484 (quoting *Dombaugh*, 20 Ohio St. at 174)) (internal quotation marks omitted).

448. *See* City of New Haven v. United States, 809 F.2d 900, 905–09 (D.C. Cir. 1987) (invalidating the Impoundment Control Act of 1974); EEOC v. CBS, Inc., 743 F.2d 969, 971–74 (2d Cir. 1984) (invalidating the Reorganization Act of 1977).

*Airlines* when they joined the Court's opinion in full in that later case.

## VII. CONCLUSION

Courts sometimes misstate the framework for severability by quoting one major case, or only one relevant passage from a major case. Unadorned citations to *Champlin* or *Alaska Airlines* will readily lead to an incomplete analysis and a faulty conclusion. Correct application of severability doctrine is increasingly important as courts and scholars intensify their focus on statutory interpretation.[449] *Free Enterprise* synthesized over a century of precedent into a comprehensive framework. Whether this proves to be a useful framework remains to be seen.

One important point to remember is that the core of severability doctrine has not changed since its inception. The only aspect of the doctrine that has ever been overruled is the early presumption of nonseverability for statutes lacking a severability clause. While this presumption has been jettisoned, the Supreme Court has not decided whether to replace it with a presumption of implied severability, or with no presumption at all. Aside from that, the focus on whether a statute's provisions are still operable and whether Congress would still have enacted the abridged statute as fulfilling Congress's negotiated bargain without the invalid provision has endured from 1876 to the present. *Champlin* did not overrule the first five decades of severability decisions. *Alaska Airlines* did not overrule *Champlin*. *Ayotte* did not overrule *Alaska Airlines*. And *Free Enterprise* did not overrule *Ayotte*. To the contrary, each major case has built upon the last.

In the first major law review Article on this issue, Robert Stern discusses his "impression left by several [severability] cases . . . that the decision on separability may have been the result of a desire to avoid more serious constitutional questions upon which the Court was divided."[450] Whether or not it originated to avoid hard choices or unpopular cases, severability has proven

449. *See* John C. Nagle, *Newt Gingrich: Dynamic Statutory Interpreter*, 143 U. PA. L. REV. 2209, 2210–11 (1995) (reviewing WILLIAM N. ESKRIDGE, JR., DYNAMIC STATUTORY INTERPRETATION (1994)).

450. Stern, *supra* note 16, at 102.

invaluable in enabling courts to fulfill their constitutional role without overstepping their bounds.

Each judicial review case must be considered on its own merits to account for many variables, and individual considerations can lead to numerous permutations. For example, take a large and complex statute in which an additional chapter was added that concerns an entirely different subject matter. Some bills are considered "must-pass" legislation, such as bills funding United States military operations. Because machinations of political factions often create impediments to Congress passing legislation, congressional leadership will sometimes attach what was originally a stand-alone piece of legislation as an amendment to a "must-pass" bill. If the resulting bill is eventually enacted as a statute, it is perfectly reasonable to imagine a scenario where a court strikes down the entire part of the statute pertaining to the invalid provision, but retains the remaining part of the legislation that was originally a separate bill.

In other circumstances, a court might not be able to firmly conclude whether such partial severability is possible. The advantage of the democratic process is that Congress can always pass a new statute if its policy objectives enjoy widespread support. So when a court invalidates an unconstitutional provision that is of major significance to the overall statutory scheme, but cannot determine exactly which provisions must fall with the defective provision, a court should strike down the entire statute to return the issue to Congress to make new legislation. This is especially true if Congress declines to include a severability clause in the statute in question.

Supreme Court decisions can shape subsequent legislation as Congress attempts to draft bills consistent with the Court's requirements to pass constitutional muster.[451] Courts faithfully applying severability doctrine could improve the quality of legislation by encouraging Congress to take care in its work. It may also encourage Congress not to develop statutes that are too large and complex, as lawmakers will be mindful that such increased size and complexity compounds the risk that the deficiency of a critical component could invalidate the entire

451. *See* J. MITCHELL PICKERILL, CONSTITUTIONAL DELIBERATION IN CONGRESS: THE IMPACT OF JUDICIAL REVIEW IN A SEPARATED SYSTEM 66–67 (2004) (citing, *e.g.*, Roe v. Wade, 410 U.S. 113 (1973)); *accord* Jona, *supra* note 197, at 714 n.101.

statute, and thus that passing shorter, discrete works where each provision is fully understood is the safer route.

As stated in the Introduction, when a statutory provision is invalidated, the question of whether that provision can be severed from the remainder turns on the significance of the provision to the statute. It is unlikely courts will abuse severability doctrine by frequently or lightly declaring an invalid provision pivotal—and thus nonseverable—to a given piece of legislation. But faithfully applying severability doctrine, and when necessary returning issues to Congress through holding a provision nonseverable, will help preserve democratic accountability and ensure that statutes on the books both conform to the Supreme Law of the Land and also enjoy the legitimacy of support by the American people.

# TROJAN HORSE: FEDERAL MANIPULATION OF STATE GOVERNMENTS AND THE SUPREME COURT'S EMERGING DOCTRINE OF FEDERALISM

MARIO LOYOLA*



* Director, Center for Tenth Amendment Studies, Texas Public Policy Foundation. B.A. European History 1995, University of Wisconsin/Madison, J.D. 1998, Washington University. I would like to thank Dr. Michael Greve and Prof. Richard Epstein for their helpful suggestions on this project. I am also deeply grateful to Brooke Rollins, Arlene Wohlgemuth, and Kathleen Hartnett White of the Texas Public Policy Foundation for their suggestions, encouragement, and unfailing support. Laura Collins and Joshua Cottle provided valuable research assistance, for which I am grateful.

## I. INTRODUCTION

It is a common refrain that the federal government has been progressively expanding its scope and reach for at least a century, regardless of which party was in control. In recent years that expansion has triggered a marked reaction among grassroots and constitutional scholars alike. In Texas, the reaction has been particularly vehement, with the state government challenging the federal administration in open defiance of its policies.[1] The reaction has focused on two policies specifically: first, the Patient Protection and Affordable Care Act (PPACA or ACA)[2], and second, the Environmental Protection Agency's (EPA) move to regulate greenhouse gases as pollutants under the Clean Air Act (CAA), which led to the partial cancellation of EPA's eighteen-year-old approval of Texas's highly successful State Implementation Plan (SIP) under the CAA.[3]

The new mandate that individuals purchase health insurance or pay a tax penalty has received the most attention of any aspect of the ACA. Texas joined twenty-five other states in successfully challenging the provision before the Eleventh Circuit, which struck down the mandate as exceeding the federal commerce power.[4] But in the Texas Legislature, another aspect of the ACA rose to the fore: namely, its provisions requiring that states expand their Medicaid rolls as a condition of continuing to receive federal Medicaid matching funds.[5]

1. *See, e.g.*, Letter from Bryan Shaw, Chairman, Tex. Comm'n on Envtl. Quality, and Greg Abbott, Att'y Gen., Texas, to Lisa Jackson, Adm'r, U.S. Envtl. Prot. Agency, and Dr. Alfredo "Al" Armendariz, Reg'l Adm'r, U.S. Envt'l Prot. Agency Region 6 (Aug. 6, 2010), http://www.tceq.texas.gov/assets/public/comm_exec/epa.pdf.

2. Pub. L. No. 111-148, 124 Stat. 119 (2010), *amended by* Health Care and Education Reconciliation Act of 2010, Pub. L. 111-152, 124 Stat. 1029 (2010). This statute is often referred to as "Obamacare."

3. Clean Air Act, 42 U.S.C. §7401 (1963); Naureen S. Malik, *EPA Rejects Texas Flexible Air-Quality Permit Authority*, WALL ST. J., June 30, 2010, http://online.wsj.com/article/SB10001424052748703426004575339140408652292.html; *Texas: Court Allows E.P.A. to Issue Greenhouse Permits*, N.Y. TIMES, Jan. 13, 2011, http://www.nytimes.com/2011/01/13/us/13brfs-COURTALLOWSE_BRF.html.

4. Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health & Human Servs., 648 F.3d 1235 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

5. Patient Protection and Affordable Care Act, 42 U.S.C. §18001 (2010); Dave Montgomery, *Conservative Legislators in Texas Seek to Opt out of Medicaid*, FORT WORTH

Facing a significant budget shortfall, and an unsustainable fiscal outlook for the state's Medicaid program (an outlook that is significantly aggravated by the ACA over the long-term), reform-minded state legislators explored every conceivable avenue for "opting out" of Medicaid entirely and replacing it with a state-based system.[6] But in the end, state legislators apparently concluded that the penalty of losing Medicaid matching funds was simply too great, and the plan went nowhere.

On the environmental front, the EPA's move to regulate greenhouse gases led to a "SIP Call" late last year.[7] A "SIP Call" is an EPA rule that prescribed the elements that a State Implementation Plan under the CAA must include in order to secure EPA approval. When a state fails to submit a conforming plan, the EPA can exact a number of penalties, including FIPing the state—that is, imposing a Federal Implementation Plan upon the state under the CAA. The EPA allows states to regulate in a federally pre-emptible area, on condition that state regulations comply with federal guidance. This is known, somewhat confusingly, as "conditional preemption."[8] The SIP Call provided that every approved SIP needs to have a provision that "automatically updates" to include any pollutant designated by the EPA.[9] The state of Texas has taken the position that the EPA's entire scheme for regulation of greenhouse gases, and in particular the "automatic update" provision of the SIP Call, violates federal law and both the federal and state constitutions. The EPA noted Texas's response, partially cancelled its eighteen-year-old approval of the state's SIP, and moved to impose a Federal Implementation Plan.[10]

---

STAR-TELEGRAM, Nov. 13, 2010, http://www.star-telegram.com/2010/11/13/2629628/conservative-legislators-in-texas.html.

6. Dave Montgomery, *Conservative Legislators in Texas Seek to Opt out of Medicaid*, FORT WORTH STAR-TELEGRAM, Nov. 13, 2010, http://www.star-telegram.com/2010/11/13/2629628/conservative-legislators-in-texas.html.

7. Action to Ensure Authority to Issue Permits under the Prevention of Significant Deterioration Program to Sources of Greenhouse Gas Emissions: Finding of Substantial Inadequacy and SIP Call, 75 Fed. Reg. 77,698 (Dec. 13, 2010) (to be codified at 40 C.F.R. pt. 52) [hereinafter "Final SIP Call"].

8. It would be more accurate to call this practice "conditional non-preemption" or "conditional permission," but in this article I will stick with the common usage among legal scholars.

9. Final SIP Call, *supra* note 7.

10. Determinations Concerning Need for Error Correction, Partial Approval and Partial Disapproval, and Federal Implementation Plan Regarding Texas Prevention of Significant Deterioration Program, 75 Fed. Reg. 82,430 (Dec. 30, 2010) (to be codified at 40 C.F.R. pt. 52) [hereinafter "Interim Partial SIP Disapproval and Interim FIP"].

Both conditional federal grants and conditional preemption insinuate the federal government deeply into the legislative and regulatory processes of state governments. Both are examples of "cooperative federalism."[11] This article argues that both practices are incompatible with "the structural framework of dual sovereignty," the standard of federalism enshrined by the Supreme Court in *Printz v. United States*.[12]

According to Dr. Michael Greve, *Printz* and related precedents have "elevate[d] the Tenth Amendment into an extra-textual, judge-made principle of intergovernmental immunity."[13] Protecting the Constitution's "structural framework of dual sovereignty" has thus emerged as a doctrine with potentially far-reaching consequences.

Because the "intergovernmental immunity" now understood to be enshrined in the Tenth Amendment is just as vulnerable to federal power indirectly applied in the guise of cooperative federalism as when such power is directly applied, there is ultimately a conflict between that "intergovernmental immunity" and the precedents that sustain conditional federal grants and conditional preemption. This article argues that as the Supreme Court examines and reexamines both practices, it should conclude (following the logic of *New York*, *Printz*, and *Bond*[14]) that neither practice can be squared with the federal structure of our Constitution, and that, in the long run, there may be no alternative to a judicially enforceable separation of federal and state government functions. Wherever federal programs confront states with a choice between subordinating local preferences to federal ones, on the one hand, and giving up either revenue or regulatory autonomy on the other, there is

11. *See, e.g.*, Ilya Somin, *Closing the Pandora's Box of Federalism: The Case for Judicial Restriction of Federal Subsidies to State Governments*, 90 GEO. L.J. 461, 465 (2002) (discussing federal conditional grants as an example of cooperative federalism), *and* New York v. United States, 505 U.S. 144, 145 (1992) (describing conditional preemption as a program of cooperative federalism).

12. Printz v. United States, 521 U.S. 898, 918, 932 (1997).

13. MICHAEL S. GREVE, REAL FEDERALISM: WHY IT MATTERS, HOW IT COULD HAPPEN 48 (1999).

14. *Cf.* New York v. United States, 505 U.S. 144 (1992) (holding that forcing states to either accept ownership of waste or regulate according to instructions of Congress is outside Congress's enumerated powers); Printz v. United States, 521 U.S. 898 (1997) (stating that obligation to conduct background checks on prospective handgun purchasers imposed unconstitutional obligation on state officers to execute federal laws); Bond v. United States, 131 S. Ct. 2355 (2011) (holding that enforcement of the Chemical Weapons Convention intruded upon police power reserved to the states).

coercion. The logic of *New York*, *Printz*, and *Bond* would not need stretching very far to reach that conclusion.

The deeper implication is the tension between individualist competition and collectivist uniformity. The tension between federalism and national majority rule is one manifestation of the great debate at the heart of modern American politics: the tension between individualism and collectivism; between those who think interstate competition is something to be protected, and those who think it is something to be protected *against*; in short, between competitive federalism and cooperative federalism.

Part II shows how the tension between competitive federalism and cooperative federalism was distilled in the law and politics of the first half of the twentieth century, as a function of the tension between federalism and nationalism. Part III traces the Supreme Court's journey away from competitive federalism to cooperative or "process" federalism and back again. Parts IV and V examine the Supreme Court jurisprudence of conditional federal grants and conditional preemption, respectively, to show in each case how the logic of those opinions is contradictory, unsustainable, and ultimately incompatible with the federal structure of our Constitution. Part VI attempts to elaborate a judicially enforceable doctrine of separation of federal and state government functions, as applied to conditional federal grants and conditional federal preemption, to serve the Supreme Court's renewed interest in defending the "structural framework of dual sovereignty."[15]

## II. THE GREAT DEBATE OF AMERICAN POLITICS: NATIONAL COLLECTIVISM V. FEDERAL INDIVIDUALISM

Though nationalist sentiment had deep cultural roots in modern Europe, the idea of national self-determination[16] as the criterion of legitimacy for a system of government can trace its birth to the Revolutions of 1848 in Europe.[17] Although that was

15. *Printz*, 521 U.S. at 932.

16. "National self-determination" is the idea that the nation should be the basic unit of sovereign political organization, and that in order to have sovereign legitimacy, a regime must be ordered so as to give expression to the will of the nation, usually as some function of national majority rule.

17. PHILIPP BOBBITT, THE SHIELD OF ACHILLES 178 (2002) ("The turning point occurred in the late 1840s, when, for similar but unrelated domestic political purposes,

when the idea took root in European political thought, it would be another seventy years before it was put fully into practice there—in the aftermath of World War I.

In the United States, national self-determination took root more gradually. Given its colonial origins, self-government in America had little use for the notion of national self-determination. At the start of the Revolution, even radicals such as John Adams claimed only that the legislatures of the various colonies were co-equal with Parliament within the Kingdom of King George III, and that Parliament therefore could not rule over the colonies.[18] The animus later turned against the King personally, during the very process of conceiving the Declaration of Independence, but even then the colonies advanced an argument of representative government that had nothing to do with national self-determination.[19] Indeed, the Declaration of Independence created thirteen "Free and Independent States" with the explicit attributes of sovereignty.[20]

Wary of the tendencies which had led other democratic experiments to end in failure, the Framers designed a constitution that went a step beyond purely national majority rule by guaranteeing majority rule at multiple levels of society, diffusing power in order to enhance both self-government and

---

European politicians seized on the idea of national self-determination as the key element underpinning a program of political reform.").

18. SAMUEL ELIOT MORISON, HENRY STEELE COMMAGER & WILLIAM E. LEUCHTENBURG, 2 THE GROWTH OF THE AMERICAN REPUBLIC 163 (7th ed. 1980) ("Independently of one another, James Wilson[,] Thomas Jefferson, and John Adams ha[d] reached the conclusion that Parliament had no rightful jurisdiction over the colonies. 'All the different members of the British Empire,' said Wilson, 'are distinct States, independent of each other, but connected together under the same sovereign in right of the same Crown.' Wilson's *Considerations on the Authority of Parliament*, Jefferson's *Summary View*, and Adams's *Novanglus* papers published this startling theory between August 1774 and February 1775. Historically they found no ground for Parliament's authority, although they admitted that the colonies had weakly accepted it; logically there was no need for it, since the colonial legislatures were competent. The colonists should honor and obey the king, follow his lead in war, observe the treaties he concluded with other princes; but otherwise govern themselves. Thus a federal solution for the problem of liberty versus authority, which John Dickinson found to be implicit in the old empire, was now made explicit by these hard-thinking Americans in 1774–75. They demanded for the Thirteen Colonies the same dominion status which Canada, Australia, New Zealand, India, Pakistan, the West Indies, and other former colonies now enjoy in the British Empire, and which is now the official basis of the British Commonwealth of Nations.")

19. *Id.* at 172–73.

20. THE DECLARATION OF INDEPENDENCE para. 5 (U.S. 1776).

individual liberty.[21] The basic idea of the Constitution was partly based on national self-determination, but contemplated a far greater degree of state and local self-rule. In fact, to the extent that national self-determination boils down to the rule of national majorities, the Constitution was designed to protect *against* any such consolidation of power, as Federalist No. 10 makes clear.[22] That is the most essential meaning of the Tenth Amendment to the Constitution, which reserves "to the States, or to the People" those powers not expressly granted to the federal government.[23]

The Tenth Amendment has been called a "truism" that adds nothing new to the constitutional scheme of limited and enumerated powers for the national government, but there is reason to doubt this view. The Tenth Amendment enshrines the concept of "reserved" powers for the states, and well into the twentieth century it was often by reference to those reserved powers that the practical limits on federal power were defined. Once the Courts began defining federal power—particularly the spending and commerce powers—by reference only to their terms in Article I, Section 8 of the Constitution, the practical limits on federal power vanished.

The nationalist projects of the Progressive Movement and Franklin Roosevelt's New Deal found one obstacle after another in the Supreme Court's interpretation of the Constitution.[24] Abetted by the political branches, popular animus against the Court increased until finally, in the 1930s, the Constitution's federalism constraints collapsed before the irredentist principle of unrestrained national majority rule. The crisis came in 1937, when, after several years in which major New Deal initiatives were struck down by the Supreme Court, FDR threatened to pack the Court with five extra justices who would vote in his

21. THE FEDERALIST NO. 10, at 47 (James Madison) (Ian Shapiro ed., 2009). *See also,* Bond v. United States, 131 S. Ct. 2355, 2364 (2011) ("Federalism secures the freedom of the individual. It allows States to respond, through the enactment of positive law, to the initiative of those who seek a voice in shaping the destiny of their own times without having to rely solely upon the political processes that control a remote central power.").

22. *Id.*

23. U.S. CONST. amend. X.

24. GREVE, *supra* note 13, 14–17.

favor.[25] The Supreme Court promptly abdicated its role as guardian of constitutional limits on federal power.[26]

That same year, Walter Lippmann published *The Good Society*, in which he argued that there are two kinds of political system, the collectivist and the individualist.[27] Among the collectivist systems he grouped Communism and Nazism, being among the first major Western intellectuals to realize that there was little difference between the two.[28] But the most striking aspect of the book was that he also grouped, among the collectivists, what he termed "gradual collectivists" or "democratic collectivists."[29]

Lippmann had been an early supporter of Woodrow Wilson and of the Roosevelts. It was Lippmann who, as an aide to Wilson, had drafted the original version of the Fourteen Points, a secular encyclical for the new faith of national self-determination.[30] But by 1937, Lippmann had soured on these nationalist excesses. He became a prominent critic of FDR's New Deal, particularly its heavy-handed disregard for the Constitution's constraints on federal power:

> The gradual collectivist believes in the absolutism of the majority, having by a fiction identified the mandates of transient majorities with the enduring and diverse purposes of the members of a community. He thinks it absurd that a few oligarchs in the Kremlin or demagogic dictators in Berlin or Rome should pretend that their personal decisions are the comprehensive purposes of great nations. Yet the gradual collectivist, under the banner of popular sovereignty, believes in the dictatorship of random aggregations of voters. In this theory the individual has no rights as against the majority, for constitutional checks and bills of rights exist only by consent of the majority.[31]

---

25. *Id.*

26. In the Warren and Burger eras, the Supreme Court embraced "legislating from the bench," overturning long-standing precedents in order to impose federal preferences in electoral apportionment, abortion rights, civil rights, and the like. In this sense, the Court went from being a guardian against unlimited federal power to a willing accomplice in its expansion.

27. WALTER LIPPMANN, THE GOOD SOCIETY 106–11 (1937).

28. *Id.*

29. *Id.*

30. John L. Snell, *Wilson on Germany and the Fourteen Points*, 26 J. MOD. HIST. 364, 365 (1954).

31. LIPPMANN, *supra* note 27, at 107.

Nearly twenty years earlier, in the case of *Hammer v. Dagenhart* (1918), the Supreme Court struck down an act of Congress prohibiting the interstate transport of goods produced in factories that had employed child labor.[32] The law did not seek to outlaw intrastate child labor directly, a subject that at the time was understood to lie entirely outside the federal power to regulate commerce "among the several States."[33] The law forbade only the *interstate transport* of goods produced in factories where child labor had been employed.[34] The majority reasoned that regardless of the actual thing regulated, the purpose and effect of the act was clearly to outlaw child labor, something Congress had no power to do, and hence, the law was unconstitutional.[35]

In a famous dissent, Justice Oliver Wendell Holmes, Jr., thought the majority's reasoning flawed becausethe purpose and consequences of the law were should be irrelevant if its subject matter lay within the federal government's enumerated powers. If Congress had the power to regulate interstate commerce, it had the power to forbid it, or any part of it, and courts had no business inquiring into Congress's purposes.[36]

Holmes ignored the danger that, through such conditions, Congress might use its interstate commerce power to coerce the states to adopt policies that had nothing to do with interstate commerce. As Prof. Richard Epstein notes, the majority "understood the statute for what it was; it was not an effort to control the goods themselves, but to prescribe the internal rules governing their manufacture within the state."[37]

Attaching conditions to the application of federal power raised obvious dangers of subverting the independence and proper functioning of state governments. Nevertheless, *Dagenhart* was eventually overruled by *United States v. Darby*.[38] Sustaining the Fair Labor Standards Act, which prohibited the interstate transport of goods produced in contravention of certain labor standards, *Darby* still maintained a formal

32. Hammer v. Dagenhart, 247 U.S. 251, 277 (1918).
33. U.S. CONST. art. I, § 8, cl. 3.
34. *Dagenhart*, 247 U.S. at 276–77.
35. *Id.* at 276.
36. *Id.* at 277–78 (Holmes, J., dissenting).
37. Richard A. Epstein, *The Proper Scope of the Commerce Power*, 73 VA. L. REV. 1387, 1427 (1987).
38. United States v. Darby, 312 U.S. 100, 116–17 (1941).

distinction between intrastate manufacturing and interstate commerce.[39] Yet he Court went further, noting that Congress's power "extends to those activities intrastate which so affect interstate commerce . . . as to make regulation of them appropriate means to the attainment of a legitimate end, the exercise of the granted power of Congress to regulate interstate commerce."[40]

So *Darby*'s repudiation of *Dagenhart* served to vindicate Holmes's insouciance over the danger that conditions attached to federal power might coerce state governments into adopting policies they didn't want and thereby subverting their representative role. .[41] *Darby*'s repudiation of *Dagenhart* also served to vindicate another argument advanced in favor of the child labor law, by the Solicitor General of the United States in that case:

> As the conviction grew that the employment of child labor was morally repugnant and socially unwise, it came to be regarded also in the light of unfair competition in trade among the States . . . . Thus, if one State desired to limit the employment of children, it was met with the objection that its manufacturers could not compete with manufacturers of a neighboring State which imposed no such limitation.[42]

Thus the "race to the bottom" argument was born into the annals of modern American political discourse, perhaps the most common justification for every new expansion of federal power. Political economists have debated whether state competition for industry and population creates more social costs than federally imposed uniformity, or whether it reduces such costs.[43] But state governments do not only, or even primarily, compete for industry; elected state officials are, of necessity, chiefly concerned with being responsive to those who put them in office. This is why many states impose even greater regulations than those called for in federal rules,[44] regardless of

39. *Id.* at 117–18.

40. *Id.* at 118.

41. *Dagenhart*, 247 U.S. at 278 (Holmes, J., dissenting).

42. *Id.* (LEXIS, found in the Syllabus before the opinion).

43. Barry Friedman, *Valuing Federalism*, 82 MINN. L. REV. 317, 408 (1997).

44. *See, e.g.*, Regional Greenhouse Gas Initiative, http://www.rggi.org (last visited Oct. 28, 2011). The Regional Greenhouse Initiative is a regional cap-and-trade system formed by ten northeastern states.

competitive consequences, and often with little discernable effect on industry and population flows.

Even more important, interstate regulatory competition is often the best way to establish the "right" level of regulation, by giving effect to local cost-benefit preferences. Even conceding that a lower level of federal uniformity creates net social costs, that is no reason to abandon the federal structure enshrined in our Constitution in favor of national majority rule unrestrained by constitutional limits. Ignoring local cost-benefit preferences, unrestrained national majority rule is bound to result in overregulation, and hence ultimately in greater social costs. Political economists have noted even more basic conceptual flaws in the "race to the bottom" argument.[45] Even in 1918, eliminating state choice was not necessary in order to eliminate child labor: *Dagenhart* did not pit child-labor states against a federal prohibition on child labor, but rather only North Carolina's twelve-year-old child labor threshold against the federal fourteen-year-old threshold.[46] By the time the case was decided, all the states had child labor laws on the books,[47] all of them equal or close to the new federal standard, and the trend was clearly in favor of universally abolishing the practice altogether. The obvious reason is that state governments were responsive to local preferences, regardless of interstate competition. There was no race-to-the-bottom. The whole argument was a figment of political advocacy.

In the year after *Darby*, the doctrine of "substantial effects" on interstate commerce, which had been percolating through Supreme Court cases (mostly in minority opinions) for several decades,[48] led to the milestone case of *Wickard v. Filburn.*[49] *Wickard* held that any activity, however local, is within the federal power to regulate commerce "among the several States" if the activity has a substantial effect on interstate commerce when *all instances of it are aggregated across the nation.*[50] Thus, even the

45. *See generally, e.g.*, Richard L. Revesz, *Rehabilitating Interstate Competition: Rethinking the "Race-to-the-Bottom" Rationale for Federal Environmental Regulation*, 67 N.Y.U. L. REV. 1210 (1992) (arguing that there is no race to the bottom over environmental standards).

46. *Dagenhart*, 247 U.S. 251.

47. GREVE, *supra* note 13, at 15.

48. THE HERITAGE FOUNDATION, THE HERITAGE GUIDE TO THE CONSTITUTION 103 (2005).

49. Wickard v. Filburn, 317 U.S. 111 (1942).

50. *Id.* at 128–29.

wheat a farmer grew for his own consumption, never to enter the stream of even local commerce, was within the federal power to regulate commerce "among the several States."[51] The aggregation principle read the clause "among the several States" straight out of the Constitution because, as Justice Thomas noted in his concurring opinion in *United States v. Lopez*, "one always can draw the circle broadly enough to cover an activity that, when taken in isolation, would not have substantial effects on commerce."[52] The aggregation principle had no stopping point: the federal government would now have the power to regulate all commerce. As a justification for unrestrained national majority rule, in the guise of federal uniformity, the "race-to-the-bottom" argument has remained firmly entrenched in our political discourse ever since. It imposes federal uniformity where the Framers intended to preserve diversity and justifies federal regulation of virtually everything the Framers assured the States' ratification conventions that the federal government would never regulate.

There is at the heart of the "race-to-the-bottom" argument an article of faith that is arguably inimical to the founding principles of our Constitution. That article of faith holds that inequality of living standards—from state to state—is a social injustice, and that is the role of government to redress that injustice by seeking federal uniformity. In this view, state regulatory competition is a force to be protected *against*; and so too the diffusion of power among multiple levels of government. Even the enhanced self-governance and individual liberty that federalism was meant to protect become drivers of social injustice in the view of the national-collectivists. The individualist emphasis on self-reliance and individual liberty, rooted in an originalist conception of the Constitution, has found itself locked in a dialectic with the national-collectivist impulse towards the largely unrestrained rule of transient national majorities. This dialectic has arguably come to define the two main approaches to federalism at the Supreme Court—and seems increasingly to define the two political parties.[53]

51. *Id.*

52. United States v. Lopez, 514 U.S. 549, 600 (1995) (Thomas, J., dissenting).

53. *See generally* ALFRED S. REGNERY, UPSTREAM 211–326 (2008). Among the national political parties, the breakdown has not been uniform. Though the Democratic party has been thoroughly collectivist and nationalist in its orientation since the administration of FDR, it was not so in the south until the 1990s. The Progressive Movement of the early

Indeed, for more than seventy years, collectivists in Congress, the White House, and the Supreme Court have worked in tandem to expand federal power dramatically in each generation.

By the 1980s, the Supreme Court had moved so far from the Constitution's constraints on federal power, that it could point to no protection for federalism save the "national political process" itself.[54] Only then did the Court begin to realize that, far from protecting federalism, the national political process is the gravest danger to it. That journey is the subject of the next section.

## III. FROM FEDERALISM TO NATIONALISM AND BACK AGAIN

In Federalist No. 45, James Madison articulated a formalistic vision of the separation of federal and state power under the proposed Constitution:

> The powers delegated by the proposed Constitution to the federal government, are few and defined. Those which are to remain in the State governments are numerous and indefinite. The former will be exercised principally on external objects, as war, peace, negotiation, and foreign commerce; with which last the power of taxation will, for the most part, be connected. The powers reserved to the several States will extend to all the objects which, in the ordinary course of affairs, concern the lives, liberties, and properties of the people, and the internal order, improvement, and prosperity of the State.[55]

Of course, Madison himself looked to the federal structure of the Constitution for the solution to a major internal problem, namely the need for "a safeguard against domestic faction and insurrection."[56] In Federalist No. 10, he fretted that purely national majority rule would render the people's representatives insensitive to local concerns, while purely local majority rule

---

twentieth century, for its part, grew out of the Republican party, which in its Northeast and Midwest "establishments" has retained a rather collectivist and nationalist orientation to this day, achieving its most expansive exponent in the ultra-nationalist administration of Richard Nixon, which expanded the scope of the federal government more than any administration before or since. The national-collectivist accumulation of power since the administration of FDR has led to a counter-reaction in the "conservative movement" identified with *National Review*, Barry Goldwater, Ronald Reagan, and, today, the Tea Party.

54. Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528, 554 (1985).

55. THE FEDERALIST NO. 45 (James Madison), *supra* note 21, at 237.

56. THE FEDERALIST NO. 10 (James Madison), *supra* note 21, at 47.

would render them insensitive to national ones, creating, in each case, a tendency toward divisive faction and even insurrection.[57] "The federal Constitution forms a happy combination in this respect; the great and aggregate interests being referred to the national, the local and particular to the State legislatures."[58]

Thus, the federal structure of the Constitution—its diffusion of power among multiple levels of government—was meant to protect against majoritarian tyrannies at every level. Integral to this conception was a formal distinction between the categories of powers granted to the federal government and those reserved for the states—the distinction enshrined in the Tenth Amendment.[59]

This conception lasted throughout the nineteenth century and well into the twentieth. In 1824, the Supreme Court held in *Gibbons v. Ogden* that navigation and commerce across state lines fall within the federal commerce power.[60] *Gibbons* rests on two pillars of the Constitution: the formal separation of federal and state functions, and the Supremacy Clause: "[T]he sovereignty of Congress, though limited to specified objects, is plenary as to those objects."[61] This was not the resounding affirmation of federal supremacy that some might suppose nowadays. Chief Justice John Marshall shared James Madison's foundational assumption that federal powers would be few and strictly defined, and that States would remain the major agents of regulation and self-government:[62]

> It is not intended to say that these words comprehend that commerce, which is completely internal, which is carried on between man and man in a State, or between different parts of the same State, and which does not extend to or affect other States. Such a power would be inconvenient, and is certainly unnecessary.

57. *Id.*

58. *Id.* at 52.

59. "The powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people." U.S. CONST. amend. X.

60. Gibbons v. Ogden, 22 U.S. 1 (1824).

61. *Id.* at 197.

62. *See id.* at 205 (discussing Congress's limited power).

> Comprehensive as the word 'among' is, it may very properly be restricted to that commerce which concerns more States than one. The phrase is not one which would probably have been selected to indicate the completely interior traffic of a State, because it is not an apt phrase for that purpose; and the enumeration of the particular classes of commerce, to which the power was to be extended, would not have been made, had the intention been to extend the power to every description.[63]

The Court observed that "inspection laws, quarantine laws, health laws of every description, as well as laws for regulating the internal commerce of a State" were but a few examples "of that immense mass of legislation" not surrendered to the federal government.[64] "No direct general power over these objects is granted to Congress," Marshall observed, "and, consequently, they remain subject to State legislation."[65] Only because it was so sure of the stringent limitations on the scope of federal power, and the preeminence of States with respect to most categories of legislation, did the Court feel so confident asserting the supremacy of federal law within its domain.[66] Hence, an expansive view of the powers reserved to the states was a necessary predicate of Marshall's expansive view of the Supremacy Clause. Otherwise, it was obvious that there would be no way to prevent that "great consolidation of Government" that Patrick Henry warned of in the Virginia ratification debates.[67]

Formal categories were indispensable to the Court's federalism jurisprudence until the New Deal, because the boundary between federal and state authority was made clearer and more stable by definition on both sides of the divide. If knowing exactly which powers had been delegated to the federal government helped us understand which powers had been reserved to the states—the idea captured in the Tenth Amendment—the reverse was also true.

Hence, when the nationalist program of FDR destroyed any tangible limit on the federal commerce power, it also destroyed the formal walls protecting the powers reserved to the States. For many decades, the Supreme Court abandoned all pretense of

63. *Id.* at 194–95.
64. *Id.* at 203.
65. *Id.*
66. *Id.*
67. JACK N. RAKOVE, ORIGINAL MEANINGS: POLITICS AND IDEAS IN THE MAKING OF THE CONSTITUTION 162 (1996).

protecting the "very extensive portion of active sovereignty"[68] retained by the states, and quickly discovered, especially under the Warren and Burger Courts, that the Justices often enjoyed legislating more than judging.[69]

During this time, federal power continued to grow along with the increasing legitimacy of national majority rule. But growing friction with core state functions was inevitable. When federalism cases once again began making their way to the Supreme Court, the justices found themselves caught in an irresolvable dilemma: How could they now defend the federal structure of the Constitution, when they themselves has long since eviscerated it?

The Court went back and forth in an embarrassing series of reversals. When the Court once again took up the Fair Labor Standards Act in *Maryland v. Wirtz*, it ruled that the Act could indeed regulate the employees of state-run schools and hospitals.[70] The Court relied heavily on this passage from another New Deal case, *United States v. California*:

> [W]e look to the activities in which the states have traditionally engaged as marking the boundary of the restriction upon the federal taxing power. But there is no such limitation upon the plenary power to regulate commerce. The state can no more deny the power if its exercise has been authorized by Congress than can an individual.[71]

Just eight years later, in *National League of Cities v. Usery*, the Court again took up the application of the Fair Labor Standards Act to state and local employees and overruled *Wirtz*:

> Congress may not exercise that power so as to force directly upon the States its choices as to how essential decisions regarding the conduct of integral governmental functions are to be made. We agree that such assertions of power, if unchecked, would indeed, as Mr. Justice Douglas cautioned in his dissent in *Wirtz*, allow "the National Government [to] devour the essentials of state sovereignty," [citations omitted] and would therefore transgress the bounds of the authority

68. THE FEDERALIST NO. 45 (James Madison), *supra* note 21, at 235.
69. REGNERY, *supra* note 53, at 211–55.
70. 392 U.S. 183, 201 (1968).
71. 297 U.S. 175, 185 (1936).

> granted Congress under the Commerce Clause .... We are therefore persuaded that Wirtz must be overruled.[72]

The commerce power did not permit Congress to infringe on state sovereignty:

> If Congress may withdraw from the States the authority to make those fundamental employment decisions upon which their systems for performance of these functions must rest, we think there would be little left of the States' "'separate and independent existence.'" Thus, even if appellants may have overestimated the effect which the Act will have upon their current levels and patterns of governmental activity, the dispositive factor is that Congress has attempted to exercise its Commerce Clause authority to prescribe minimum wages and maximum hours to be paid by the States in their capacities as sovereign governments. In so doing, Congress has sought to wield its power in a fashion that would impair the States' "ability to function effectively in a federal system."[73]

A subsequent case, *Hodel v. Virginia Surface and Mining Reclamation Association,* elaborated the rule of *National League of Cities* into a three part-test.[74] In order for a federal law to infringe impermissibly on state sovereignty it had to: (1) regulate the "States as States"; (2) address matters that are "attributes of State sovereignty"; and (3) impair state operations in their "traditional governmental functions."[75]

The three-part balancing test articulated in *Hodel* was already arguably removed from *National League of Cities*'s bright-line defense of those aspects of state sovereignty deemed essential "to the States' separate and independent existence" and their "ability to function effectively in a federal system."[76] But the focus now shifted to the third *Hodel* requirement, namely whether the congressional exercise of the commerce power infringed on "traditional governmental functions."

---

72. Nat'l League of Cities v. Usery, 426 U.S. 833, 855 (1976).

73. *Id.* at 851–52 (internal citations omitted).

74. 452 U.S. 264 (1981). This was later elaborated into a four-part test that included the added requirement that it could not be a case in which the federal interest "justifies state submission." *See* Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528, 537 (1985).

75. *Hodel,* 452 U.S. at 287–88, n.29.

76. *Nat'l League of Cities,* 426 U.S. at 843, 845.

The Supreme Court soon reversed itself again.[77] In 1984, the Court again took up the Fair Labor Standards Act in *Garcia v. San Antonio Metropolitan Transit Authority*, this time as applied to employees of the local transit authority.[78] The Court surveyed the landscape of federal court decisions trying to apply the "traditional governmental functions" test and found an incomprehensible cacophony of rulings. "We find it difficult, if not impossible, to identify an organizing principle that places each of the cases in the first group on one side of a line and each of the cases in the second group on the other side."[79] The Court essentially revived the ruling in *Wirtz*: "We therefore now reject, as unsound in principle and unworkable in practice, a rule of state immunity from federal regulation that turns on a judicial appraisal of whether a particular governmental function is 'integral' or 'traditional.'"[80]

Brazenly rewriting constitutional history, the Court now decided that:

> In short, the Framers chose to rely on a federal system in which special restraints on federal power over the States inhered principally in the workings of the National Government itself, rather than in discrete limitations on the objects of federal authority. State sovereign interests, then, are more properly protected by procedural safeguards inherent in the structure of the federal system than by judicially created limitations on federal power.[81]

Justice O'Connor's dissent noted the ominous implications in the majority's embrace of national majority rule as the sole guarantor of constitutional federalism: "With the abandonment of *National League of Cities*, all that stands between the remaining essentials of state sovereignty and Congress is the latter's underdeveloped capacity for self-restraint."[82] Joining her in dissent, Justice Rehnquist was simply exasperated: "I do not think it incumbent on those of us in dissent to spell out further the fine points of a principle that will, I am confident, in time again command the support of a majority of this Court."[83]

77. *Garcia*, 469 U.S. 528.
78. *Id.* at 531.
79. *Id.* at 539.
80. *Id.* at 546–47.
81. *Id.* at 552.
82. *Id.* at 588 (O'Connor, J., dissenting).
83. *Id.* at 580 (Rehnquist, J., dissenting).

He was right: *Garcia* would prove the low point in the Court's prostration before national majority rule and was soon overruled tacitly if not yet expressly. In 1992, there emerged *New York v. United States*,[84] the first of a line of cases that would establish a new bright-line rule: the federal government cannot compel a state government to do anything.[85]

This time, Justice O'Connor wrote for the majority: "While Congress has substantial powers to govern the Nation directly, including in areas of intimate concern to the States, the Constitution has never been understood to confer upon Congress the ability to require the States to govern according to Congress' instructions."[86] The Court struck down the "take title" provisions of the Low-Level Radioactive Waste Policy Amendments Act because it required states either to take title to low-level radioactive waste generated within their borders, or regulate its disposal according to Congress's instruction.[87] "In this provision," reasoned the majority, "Congress has crossed the line distinguishing encouragement from coercion."[88] Congress could not force states to choose between two alternatives neither of which Congress had the power to impose "as a free standing requirement."[89]

The decision was justly well-received in federalism circles, but it was not without its problems. The Court reaffirmed the legitimacy of both conditional federal grants and conditional preemption as forms of "encouragement" not rising to the level of "coercion."[90] The Court noted, "[w]here Congress encourages state regulation rather than compelling it, state governments remain responsive to the local electorate's preferences; state officials remain accountable to the people."[91] On the other hand, "[a]ccountability is . . . diminished when, due to federal coercion, elected state officials cannot regulate in accordance

84. 505 U.S. 144 (1992).

85. *See generally* Printz v. United States, 521 U.S. 898 (1997) (striking down a federal statute directing the operations of state officials); New York v. United States, 505 U.S. 144 (1992) (striking down a federal mandate on state nuclear waste disposal).

86. *New York*, 505 U.S. at 162.

87. *Id.* at 174–75.

88. *Id.*

89. *Id.*

90. *Id.* at 173–74 (citing Hodel v. Va. Surface Mining & Reclamation Ass'n, Inc., 452 U.S. 264 (1981) and FERC v. Mississippi, 456 U.S. 742, 764–65 (1982)).

91. *Id.* at 168.

with the views of the local electorate in matters not pre-empted by federal regulation."[92]

The weakness in the Court's reasoning was that, given the broad sweep of the federal tax-and-spend power (in the case of conditional federal grants) and of federal commerce power (in the case of conditional preemption), it does not take much "encouragement" to diminish a state government's responsiveness to local preferences. By definition any such "encouragement" diminishes a state government's responsiveness to local preferences in favor of national ones, the only variable being a matter of degree, that is the whole purpose of such "encouragement." If we reverse the logic of O'Connor's distinction between encouragement and coercion, and start by asking whether a federal law leaves elected state officials free to regulate "in accordance with the views of the local electorate,"[93] it becomes obvious that virtually all cases of federal "encouragement" boil down to some degree of coercion. That is the subject of the next two sections of this article.

Before pursuing the Court's reasoning into the realm of conditional federal grants and conditional coercion, two more cases bear mentioning, including the most important and far-reaching of the Court's commandeering cases, *Printz v. United States.*[94]

In *Printz,* the Court struck down a part of the Brady Act that required states to conduct background checks on prospective gun purchasers.[95] The Court ruled that because the federal government cannot compel state governments to regulate, neither can it compel state officials to perform any particular function.[96] The ruling was a welcome relief from the modern plague of indeterminate multi-prong balancing tests. Such tests, wrote Justice Scalia for the majority,

> might be relevant if we were evaluating whether the incidental application to the States of a federal law of general applicability excessively interfered with the functioning of state governments. [Citations omitted] But where, as here, it is the whole *object* of the law to direct the functioning of the state

92. *Id.* at 170.
93. *Id.* at 169.
94. 521 U.S. 898 (1997).
95. *Id.* at 935.
96. *Id.*

> executive, and hence to compromise the structural framework of dual sovereignty, such a "balancing" analysis is inappropriate. It is the *very principle* of separate state sovereignty that such a law offends, and no comparative assessment of the various interests can overcome that fundamental defect.[97]

The ruling in *Printz* was categorical. The federal government could not use the commerce power to compel state officials to perform any function,[98] period. It did not matter if the function was minor or the federal interest overwhelming. It did not matter if it was a traditional state function or not. It did not matter if the state incurred no costs at all. If a federal law offended "the structural framework of dual sovereignty"[99] it was now categorically unconstitutional.

The most potentially consequential feature of the opinion in *Printz* was the revival of the notion that the federal and state governments occupy separate spheres in a "structural framework of dual sovereignty" and that the States must remain "independent and autonomous within their proper sphere of authority."[100] The protection for this federal structure was further reinforced by Scalia's invocation of the Necessary and Proper Clause.[101] A law which violates the federal structure of the Constitution is not a law that is "proper" for carrying into execution any enumerated power, "and is thus, in the words of the Federalist, 'merely [an] ac[t] of usurpation' which 'deserves to be treated as such.'"[102]

The Court's reasoning in *Printz* could have enormous implications. If states must remain "independent and autonomous within their proper sphere of authority,"[103] and any law which crosses into the sphere is not "proper" within the Necessary and Proper Clause, then it may once again be possible to trace the outer boundaries of the federal government's delegated powers by tracing the outer boundaries of the states' reserved powers.

---

97. *Id.* at 932 (citations omitted).

98. *Id.* at 935.

99. *Id.* at 932.

100. *Id.* at 928.

101. *Id.* at 923–24.

102. *Id.* (alteration in original) (quoting THE FEDERALIST NO. 33, at 204 (Alexander Hamilton)).

103. *Id.* at 924.

This view of the federal structure of the Constitution was reaffirmed last summer in *United States v. Bond.*[104] Holding that citizens have standing to challenge federal violations of state sovereignty, Justice Kennedy, for the majority, reiterated that "'federalism secures to citizens the liberties that derive from the diffusion of sovereign power.'"[105] Those laws issued by the separate spheres of authority "in excess of delegated governmental power cannot direct or control"[106] the actions of the individual. Where the federal government encroaches upon the political territory of the States, the "individual has a direct interest in objecting to laws that upset the constitutional balance" because holding true to "federalism is not for the States alone to vindicate."[107] In short, the separation of government authority into two bodies of government "protects the liberty of the individual from arbitrary power."[108] For the first time since the New Deal, the Court had finally, however unwittingly, equated unbridled national majority rule with "arbitrary power."

## IV. CONDITIONAL FEDERAL GRANTS

In *United States v. Butler,* the Supreme Court noted that through the device of conditional federal grants, "constitutional guarantees, so carefully safeguarded against direct assault, are open to destruction by the indirect, but no less effective, process of requiring a surrender, which, though, in form voluntary, in fact lacks none of the elements of compulsion."[109] This indispensable observation has lied largely unnoticed in the chaff of Supreme Court dicta for decades, to virtually no effect. Nowadays, those seeking protection from the dictates of conditional federal funds must look to the woefully insufficient standard of *South Dakota v. Dole.*[110]

The ACA's Medicaid expansion provisions show how illusory state "prerogative" really is in the conditional federal grants context. The federal grant constitutes forty percent of all federal funds paid to States.[111] It is nearly impossible to imagine that any

104. Bond v. United States, 131 S. Ct. 2355 (2011).
105. *Id.* at 2364 (quoting New York v. United States, 505 U.S. 144, 181 (1992)).
106. *Id.*
107. *Id.*
108. *Id.*
109. 297 U.S. 1, 72 (1936).
110. 483 U.S. 203 (1987).
111. U.S. CENSUS BUREAU, FEDERAL AID TO STATES FOR FISCAL YEAR 2009 vii (2010).

state government would find it politically feasible to forego such a large amount of its citizen's taxes paid to the federal government for any reason, and none ever has.[112]

*Dole* upheld a federal law that threatened states with the loss of five percent of federal highways funds if they did not raise their drinking age to twenty-one.[113] The Court noted that the penalties attaching to such conditional federal programs could not be so onerous as to pass "the point at which pressure turns into compulsion."[114] *Dole* insists that state prerogative must be preserved, both in theory and in fact, but would have us believe that freedom of choice is preserved in the state's ability to refuse the funding and its conditions.

But any amount of money taxed away from the states and returned to them only on condition of compliance with federal preferences weakens the state's ability to choose. The only question is whether it weakens that freedom of choice a little or a lot— a question not of kind but of degree. *Dole* conflates the sliding scale of coercion with an imaginary spectrum along which pressure is supposed turn into compulsion at some point. But this is a logical fallacy. If the penalty involved is miniscule, there is pressure, and freedom of choice is lessened; if the penalty is enormous, there is still freedom of choice, notwithstanding the pressure. Either there is coercion in both cases or there is coercion in neither. *Dole*'s attempt to articulate some way of distinguishing between compulsion and mere encouragement was doomed to be unworkable in practice, and so it has proved. The *Dole* standard has never triggered a ruling of coercion, no matter how great the penalty.[115]

The coercion problem is particularly acute where the federal government makes more onerous the conditions attaching to an existing program in which the States are already heavily invested. That, in a nutshell, is what the Medicaid expansion

112. JAGADEESH GOKHALE, CATO INSTITUTE, THE NEW HEALTH CARE LAW'S EFFECT ON STATE MEDICAID SPENDING 6 (2011), http://www.cato.org/pubs/wtpapers/StateMedicaidSpendingWP.pdf.

113. *Id.* at 205.

114. South Dakota v. Dole, 483 U.S. 203, 211 (1987) (quoting Steward Mach. Co. v. Davis, 301 U.S. 548, 590 (1937)).

115. *See* Florida *ex rel.* Bondi v. U.S. Dep't of Health & Human Servs., 780 F. Supp. 2d 1256, 1266–69 (N.D. Fla. 2011) (discussing *Dole*), *aff'd in part and rev'd in part*, 648 F.3d 1235, 1328 (11th Cir. 2011) (affirming the judgment regarding two provisions but reversing with regard to severability), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

provisions of the ACA do. If *Dole* was ever going to be used to establish the coercive effect of a federal conditional grant program, *HHS v. Florida*—the main challenge to the ACA—was the textbook case.[116] Both the district court and the Eleventh Circuit Court of Appeals struck down the individual mandate in ACA, on federalism grounds.[117] But both refused to find coercion in the ACA's Medicaid expansion provisions.[118]

In attempting to trace the limits on the federal conditional spending power, the Supreme Court in *Dole* observed, "[t]he spending power is of course not unlimited, but is instead subject to several general restrictions articulated in our cases."[119] The Court listed four: (1) the exercise of the spending power must be in pursuit of "the general welfare"; (2) the conditions must be unambiguously stated; (3) the conditions must be related to the federal interest in particular national projects or programs; and (4) the conditions cannot require the States to do something that is otherwise unconstitutional.[120]

Each of these limits either by its own terms offers no logical protection for state sovereignty, or has been applied by federal courts in a way that offers no protection. First, in applying the "general welfare" prong, federal courts must "defer substantially" to the judgment of Congress,[121] and the Court has even speculated that the standard is not judicially enforceable at all.[122] The second restriction, that conditions be unambiguously stated, is an issue of statutory drafting with no bearing on the nature or scope of the condition, or whether it constitutes coercion. The third restriction, that the condition bear a reasonable relation to the federal interest in a national project or program, has the promise implied in Justice Sandra Day O'Connor's dissent in *Dole*, namely that of drawing a distinction between conditions on how the federal grant is to be spent (which O'Connor thought permissible) and conditions based on state adoption of a regulatory scheme not specifically related to how the grant is to be spent (which O'Connor thought

116. *See generally Florida* ex rel. *Bondi*, 780 F. Supp. 2d 1256.

117. *Florida* ex rel. *Att'y Gen.*, 648 F.3d at1302–07; *Florida* ex rel. *Bondi.*,780 F. Supp. at 1298.

118. *Id.*

119. South Dakota v. Dole, 483 U.S. 203, 207 (1987).

120. *Id.* at 207–08.

121. *Id.*

122. *Id.* at 207 & n.2.

impermissible).[123] But the Court's holding in *Dole* forecloses that promise as a useful distinction, because the drinking-age requirement at issue in *Dole* was not a condition on how the federal highway funds were to be spent, but rather only a loosely related regulation.[124] And in any case *Dole* implicitly recognized that the conditions attaching to federal conditional funds may "pass the point at which pressure turns into compulsion" even if the conditions are focused purely on how the funds are to be spent.[125] The fourth restriction, the bar against requiring states to do anything that is otherwise unconstitutional, is logically of no help because we are questioning the imposition of federal conditions on state regulatory powers that we presuppose to be constitutional.

The supposed "coercion doctrine" lies in *Dole*'s observation that the penalty of losing federal funds "might be so coercive as to pass the point at which pressure turns into compulsion,"[126] which is unconstitutional. The Court insisted that compliance with federal conditions must remain "the prerogative of the States not merely in theory but in fact."[127]

*Dole* focused on the fact that unwilling states stood to lose "a relatively small percentage of certain federal highway funds":[128]

> When we consider, for a moment, that all South Dakota would lose if she adheres to her chosen course as to a suitable minimum drinking age is 5% of the funds otherwise obtainable under specified highway grant programs, the argument as to coercion is shown to be more rhetoric than fact.[129]

This mere "temptation" in the form of "relatively mild encouragement" was not enough to rise to the level of coercion.[130] According to the Court, regulatory authority over the State's drinking age "remains the prerogative of the States not merely in theory but in fact."[131]

123. *Id.* at 216.
124. *Id.* at 211.
125. *Id.* (citing Steward Mach. Co. v. Davis, 301 U.S. 548, 590 (1937)).
126. *Dole*, 483 U.S. at 211.
127. *Id.* at 211–12.
128. *Id.* at 211.
129. *Id.*
130. *Id.*
131. *Id.*

This has to be read together with preceding quotation from *Steward Machine Company*, in which the Court observed, "[b]ut to hold that motive or temptation is equivalent to coercion is to plunge the law into endless difficulties. The outcome of such a doctrine is the acceptance of a philosophical determinism by which choice becomes impossible."[132] Temptation, then, is not the same as coercion; when the States can "theoretically" opt out of a federal program, doing so remains their prerogative. But *Dole* also noted that the States must be able to retain their prerogatives "in fact."[133]

This standard raises impossible conceptual problems, which is why no federal conditional grant program has ever been found to be coercive. The incoherence of the Eleventh Circuit's ruling on the Medicaid expansion provisions of the ACA, in *Florida* ex rel. *Attorney General v. United States Department of Health and Human Services* shows why.[134] In the trial court below, Judge Roger Vinson ruled that the law's Medicaid provisions are constitutional.[135] He observed that federal courts routinely pay lip service to *Dole*'s coercion doctrine but have never in practice found coercion in any case, no matter how onerous the conditions.[136] He ruled in effect that there is no doctrine of coercion, and concluded that because the plaintiffs' coercion claim could not succeed no matter how large the penalty in fact, "the defendants are entitled to judgment as a matter of law."[137]

The twenty-six states challenging the ACA argued that they simply could not afford the loss of Medicaid funds, so compliance is in no sense voluntary.[138] If true, that would violate *Dole*. Even if opting out remained a state prerogative in theory,

132. Steward Mach. Co. v. Davis, 301 U.S. 548, 589–90 (1937).

133. *Dole*, 483 U.S. at 212.

134. Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health & Human Servs., 648 F.3d 1235 (11th Cir. 2011), *cert. granted sub nom.* Nat'l Fed'n of Indep. Business v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

135. Florida *ex rel.* Bondi v. U.S. Dep't of Health & Human Servs., 780 F. Supp. 2d 1256 (N.D. Fla. 2011), *aff'd in part and rev'd in part*, 648 F.3d 1235, 1328 (11th Cir. 2011) (affirming the judgment regarding two provisions but reversing with regard to severability), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

136. *See id.* at 1268 ("[E]very single federal Court of Appeals called upon to consider the issue has rejected the coercion theory as a viable claim.").

137. *Id.* at 1269.

138. *See id.* at 1269 (discussing that because Medicaid is the single largest federal grant-in-aid program to the states, the states effectively have no choice other than to participate in the program).

the political process itself virtually guaranteed that it could not remain so in fact: the penalty was simply too large,[139] and the federal matching funds were paid for by state residents to begin with. The government countered with evidence that in fact the penalty is less onerous than claimed.[140]

Vinson noted: "In short, there are numerous genuine disputed issues of material fact with respect to this claim that cannot be resolved on summary judgment."[141] But he nevertheless ruled that given the failure of federal courts to develop any applicable coercion standard, there really was no issue of material fact, and the government was entitled to judgment as a matter of law.[142] This was tantamount to holding that the *Dole* standard doesn't even exist. At the very least, the Medicaid count should have proceeded to a trial on the facts. *Dole* seems to require a factual inquiry into whether federal conditions "pass the point at which pressure turns into compulsion[;]"[143] compliance must remain a state prerogative "not merely in theory but in fact."[144] There was at least an issue of material fact as to whether the Medicaid expansion provisions are so onerous that states can't afford to opt out.

Hence the Eleventh Circuit should have reversed that summary judgment and returned the case to Judge Vinson for a trial on the facts. Instead it affirmed his judgment, but virtually ignored what he actually said:

> If anything can be said of the coercion doctrine in the Spending Clause context, however, it is that it is an amorphous one, honest in theory but complicated in application. But this does not mean that we can cast aside our duty to apply it;

139. *See* TEX. HEALTH AND HUMAN SERVS. COMM'N & TEX. DEP'T OF INS., IMPACT ON TEXAS IF MEDICAID IS ELIMINATED, H. 81 (2009).

140. *Id.* at 1267 ("In their voluminous materials filed in support of their motion for summary judgment, the state plaintiffs have identified some serious financial and practical problems that they are facing under the Act, especially its costs. They present a bleak fiscal picture. At the same time, much of those facts have been disputed by the defendants in their equally voluminous filings; and also by some of the states appearing in the case as *amici curiae*, who have asserted that the Act will in the long run save money for the states. It is simply impossible to resolve this factual dispute now as both sides' financial data are based on economic assumptions, estimates, and projections many years out. In short, there are numerous genuine disputed issues of material fact with respect to this claim that cannot be resolved on summary judgment.").

141. *Id.*

142. *Id.* at 1269.

143. *Id.* at 1266 (internal citations omitted).

144. South Dakota v. Dole, 483 U.S. 203, 211–12 (1987).

> indeed, it is a mystery to us why so many of our sister circuits have done so. To say that the coercion doctrine is not viable or does not exist is to ignore Supreme Court precedent, an exercise this Court will not do . . . . If the government is correct that Congress *should* be able to place any and all conditions it wants on the money it gives to the states, then the Supreme Court must be the one to say it.[145]

But by affirming Judge Vinson's summary judgment, the Eleventh Circuit in effect said that Congress should be able to place any and all conditions it wants on the money it gives to the states. It should be no mystery why so many "sister circuits" have tossed *Dole*'s coercion doctrine aside. Coercion is coercion, whether it's a single dollar or a million. A conditional federal grant is categorically coercive "because it necessarily conditions the exercise of one right upon the conscious surrender of a second."[146]

The right a state surrenders when accepting federal conditions is, of course, the right to be responsive to local preferences. The right it surrenders when refusing the federal conditions is the right to share in the benefits of a program its citizens are paying for. In essence every federal conditional grant boils down to this: Accept national majority rule within the sphere of traditional state authority, or suffer the massive transfer of funds from your state to states that do accept national preferences. This certainly qualifies as "encouragement." It is also coercion.

The *Dole* test is worse than a mirage. It does nothing tangible to protect states' regulatory autonomy, while allowing the federal courts to pretend that the states are protected. In fact, only budgetary constraints prevent the federal government from using conditional grants for the total subversion of state governments. *Garcia* stood for the proposition that the national political process is enough to protect federalism. But if, after *New York* and *Printz*, *Garcia* is no longer good law, where does that leave *Dole*?[147]

145. Florida *ex rel.* Att'y Gen. v. U.S. Dep't of Health & Human Servs., 648 F.3d 1235, 1266–67 (11th Cir. 2011) (emphasis in original), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

146. Richard A. Epstein and Mario Loyola, *ObamaCare's Next Constitutional Challenge*, WALL ST. J. (June 7, 2011), http://online.wsj.com/article/SB10001424052702304474804576367690213892556.html.

147. See Section III for discussion of *Garcia*, *New York*, and *Printz*.

In *Printz* the Court held that where a federal action threatens the "dual sovereignty" guaranteed to the states, it offends the federal structure of the Constitution, and must be struck down.[148] Judge Vinson acknowledged this, and all but invited the Supreme Court to overrule *Dole* and extend the logic of *Printz* to the arena of conditional federal grants:

> Some have suggested that, in the interest of federalism, the Supreme Court should revisit and reconsider its Spending Clause cases [e.g., Dole]. *See* Lynn A. Baker, *The Spending Power and the Federalist Revival*, 4 Chap. L. Rev. 195–96 (2001) (maintaining the "greatest threat to state autonomy is, and has long been, Congress's spending power" and "the states will be at the mercy of Congress so long as there are no meaningful limits on its spending power"). However, unless and until that happens, the states have little recourse to remaining the very junior partner in this partnership.[149]

Judge Vinson reasoned that the *Dole* coercion standard doesn't really exist, and in effect, all federal conditional grant programs are permissible, no matter how great the penalty. With glaring incoherence, the Eleventh Circuit affirmed his ruling while flatly contradicting his basis for it. The federal courts have indeed applied the *Dole* standard in a way that permits all conditional grants. Nothing except *Garcia*'s national political process stands in the way of the total absorption of state budgets by the federal government.

Though the Commerce Clause portion of *United States v. Butler* is no longer valid law,[150] there is presumably no argument with *Butler*'s observation on the inherent conflict between a too-expansive reading of the Spending Clause, and the federal structure of the Constitution:

> Hamilton himself, the leading advocate of broad interpretation of the power to tax and to appropriate for the general welfare, never suggested that any power granted by the Constitution could be used for the destruction of local self-

148. 521 U.S. 898 (1997).

149. Florida *ex rel.* Bondi v. U.S. Dep't of Health & Human Servs., 780 F. Supp. 2d 1256, 1269 (N.D. Fla. 2011), *aff'd in part and rev'd in part*, 648 F.3d 1235, 1328 (11th Cir. 2011) (affirming the judgment regarding two provisions but reversing with regard to severability), *cert. granted sub nom.* Nat'l Fed'n of Indep. Bus. v. Sebelius, 80 U.S.L.W. 3198 (Nov. 14, 2011) (Nos. 11-393, 11-398, 11-400).

150. Wickard v. Filburn, 317 U.S. 111 (1942) (overruling *U.S. v. Butler* as to Commerce Clause standard).

> government in the states. Story countenances no such doctrine. It seems never to have occurred to them, or to those who have agreed with them, that the general welfare of the United States (which has aptly been termed 'an indestructible Union, composed of indestructible States,') might be served by obliterating the constituent members of the Union. But to this fatal conclusion the doctrine contended for would inevitably lead. And its sole premise is that, though the makers of the Constitution, in erecting the federal government, intended sedulously to limit and define its powers, so as to reserve to the states and the people sovereign power, to be wielded by the states and their citizens and not to be invaded by the United States, they nevertheless by a single clause gave power to the Congress to tear down the barriers, to invade the states' jurisdiction, and to become a parliament of the whole people, subject to no restrictions save such as are self-imposed. The argument, when seen in its true character and in the light of its inevitable results, must be rejected.[151]

If conditional federal grants have no limit in a doctrine of coercion, then the exercise of the power to make them has no limit, save such as may be self-imposed by the federal government. The failure of the federal courts to fashion the *Dole* coercion doctrine into any meaningful limit on the federal government's ability to subvert state governments through conditional grants should sooner or later tempt the Court to see the same danger in conditional grants that it has found in commandeering. The federal uniformity such grants are designed to achieve have consistently come at the expense of state governments' responsiveness to local preferences and accountability for government policies at all levels.

## V. CONDITIONAL PREEMPTION

At first blush the practice of giving states permission to regulate in a pre-emptible field, on condition that they meet federal guidelines, seems to present an entirely different problem than conditional federal grants. Conditional federal grants allow the federal government to get indirectly something it cannot get directly, namely specific state legislation.[152]

151. United States v. Butler, 297 U.S. 1, 77–78 (1936).

152. Ruth Mason, *Federalism and the Taxing Power*, 99 CALIF. L. REV. 975, 983 (2011) (discussing Congressional authority to use conditional grants to induce state legislatures to enact certain legislation).

Conditional preemption, by contrast, allows the federal government to get indirectly something that it *can* get directly, namely the regulation of individuals subject to overlapping state and federal authority.[153] The greater power of wholesale preemption, we are told, includes the lesser power of conditional preemption, and so there is not even a theoretical possibility of coercion.[154]

But, because conditional preemption accomplishes federal ends through the instruments of state government, the same danger is raised as in the conditional federal grant cases, namely, that the "inducement offered by Congress might . . . pass the point at which pressure [on state governments] turns into compulsion."[155] As with the ACA's Medicaid expansion provisions, instances of conditional preemption demonstrate the lack of meaningful difference in *New York*'s supposed distinction between "encouragement" and "coercion," and raise every bit as much concern for the "separate and independent existence" of the states.[156] Conditional preemption presents state governments with a "choice" that is really no choice at all: Either regulate this or that area according to federal preferences, or the federal government will preempt your ability to regulate it at all.

EPA's fantastical voyage into the realm of planetary climate control is a good example of how little choice in fact conditional preemption programs leave to the states. The Clean Air Act (CAA) was designed to regulate emissions of pollutants that cause direct harm to human health.[157] As Justice Scalia noted in his dissent in *Massachusetts v. EPA*, "regulating the buildup of CO2 and other greenhouse gases in the upper reaches of the atmosphere, which is alleged to be causing global climate change, is not akin to regulating the concentration of some substance that is *polluting the air*."[158] In order to bring greenhouse gases within the CAA, the EPA had to devise a

153. *See* Somin, *supra* note 11 and accompanying text.

154. Ronald D. Rotunda, *The Doctrine of Conditional Preemption and Other Limitations on Tenth Amendment Restrictions*, 132 U. PA. L. REV. 289, 324 n.250 (1984) (using the example of Congress's power to preempt all highway operations as justification for Congress's power to require states to pay their highway commissioners and toll-takers a federal minimum wage).

155. South Dakota v. Dole, 483 U.S. 203, 211 (1987).

156. *See* New York v. United States, 505 U.S. 144, 175 (1992).

157. 42 U.S.C. § 7401(b)(1).

158. 549 U.S. 497, 559 (2007).

fiendishly convoluted series of rule-makings.[159] Texas has joined other states in fighting all of these new rules in federal court.[160]

On June 3, 2010, the EPA gave the states until August 2, 2010 to report whether their SIPs would include greenhouse gases as pollutants "subject to regulation" under the CAA.[161] The government of Texas replied that it had "neither the authority nor the intention" of complying with the EPA's new greenhouse gas rules, which it considered to be illegal and unconstitutional.[162] On December 13, 2010, the EPA issued a SIP Call for thirteen states, including Texas, indicating that to secure or maintain EPA approval, SIPs would henceforth need to "automatically update" to cover all substances designated as pollutants by the EPA now or in the future.[163] In comments to the SIP Call, the Texas Commission on Environmental Quality (TCEQ) reiterated its position that the EPA's demand for a revision to the Texas SIP was illegal.[164] In response, the EPA subsequently issued a partial disapproval of the Texas SIP.[165] It explained that its original approval of the Texas SIP, eighteen

159. Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act, 74 Fed. Reg. 66,496 (Dec. 15, 2009) (to be codified at 40 C.F.R. ch. I) [hereinafter "Endangerment Finding"]; Reconsideration of Interpretation of Regulations That Determine Pollutants Covered by Clean Air Act Permitting Programs, 75 Fed. Reg. 17,004 (Apr. 2, 2010) (to be codified at 40 C.F.R. pts. 50, 51, 70, 71) [hereinafter "Timing Rule"]; Light-Duty Vehicle Greenhouse Gas Emission Standards and Corporate Average Fuel Economy Standards; Final Rule, 75 Fed. Reg. 25,324 (May 7, 2010) (to be codified at 40 C.F.R. pts. 85, 86, 600) [hereinafter "Tailpipe Rule"]; Prevention of Significant Deterioration and Title V Greenhouse Gas Tailoring Rule, 75 Fed. Reg. 31,514 (June 3, 2010) (to be codified at 40 C.F.R. pts. 51, 52, 70, 71) [hereinafter "Tailoring Rule"].

160. Determinations Concerning Need for Error Correction, Partial Approval and Partial Disapproval, and Federal Implementation Plan Regarding Texas Prevention of Significant Deterioration Program, 75 Fed. Reg. 82,430 (Dec. 30, 2010) (to be codified at 40 C.F.R. pt. 52).

161. *See* Tailoring Rule, *supra* note 159.

162. *See* Letter, *supra* note 1.

163. Action to Ensure Authority to Issue Permits under the Prevention of Significant Deterioration Program to Sources of Greenhouse Gas Emissions: Finding of Substantial Inadequacy and SIP Call, 75 Fed. Reg. 77,698 (Dec. 13, 2010) (to be codified at 40 C.F.R. pt. 52) [hereinafter "Final SIP Call"].

164. Texas Commission on Environmental Quality Comments on Actions to Ensure Authority to Issue Permits Under the Prevention of Significant Deterioration Program to Sources of Greenhouse Gas Emissions; Finding of Substantial Inadequacy and SIP Call, Docket ID No. EPA-HQ-OAR-2010-0107, FRL-9190-7 Federal Implementation Plan (FIP), Docket ID. No. EPA-HQ-OAR-2010-0107, FRL-9190-8 (Oct. 4, 2010) [hereinafter TCEQ comments].

165. Determinations Concerning Need for Error Correction, Partial Approval and Partial Disapproval, and Federal Implementation Plan Regarding Texas Prevention of Significant Deterioration Program, 75 Fed. Reg. 82,430 (Dec. 30, 2010) (to be codified at 40 C.F.R. pt. 52) [hereinafter "Interim Partial SIP Disapproval and Interim FIP"].

years earlier, had been an error because the SIP did not automatically update for newly designated pollutants.[166] In the same rulemaking, it immediately imposed a FIP as an interim final rule[167] and thereby took over permitting authority in Texas for greenhouse gas emissions under the CAA's provisions on Prevention of Significant Deterioration (PSD). As a consequence, Texas companies wishing to build new facilities or expand existing ones may now have to file for two separate permits, one covering greenhouse gas emissions, to be issued directly by the EPA, and the other covering emissions of "conventional" (i.e., actual) pollutants, to be issued as before by the TCEQ.[168] If Texas does not submit a revised SIP by December 1, 2011, the EPA could move towards a full takeover of PSD permitting authority.[169]

Texas's court challenges focus on the host of statutory questions raised by the EPA's actions, but it has not argued that the SIP Call and subsequent FIP are unconstitutionally coercive or commandeering.[170] That is understandable, given standing precedents of the Supreme Court. Its two major conditional preemption cases—*Hodel* and *FERC v. Mississippi*—both appear to foreclose any constitutional challenge on federalism grounds.[171] However, as has been noted, the Court's decisions in this area have been unstable and contradictory for decades; raising legitimate questions about how settled this area of the law really is, even after *Printz*.[172]

*Hodel* and *FERC* both relied on *National League of Cities*'s elevation of the Tenth Amendment to a principle of immunity for the "traditional governmental functions" of states.[173] Both

166. *Id.* at 82,432–33.

167. *Id.* at 82,448. The interim actions were finalized in Determinations Concerning Need for Error Correction, Partial Approval and Partial Disapproval, and Federal Implementation Plan Regarding Texas's Prevention of Significant Deterioration Program, 76 Fed. Reg. 25,178 (May 3, 2011) (to be codified at 40 C.F.R. pt. 52) [hereinafter "Final Partial SIP Disapproval and Final FIP"].

168. *See* Tailoring Rule, *supra* note 159.

169. *See* TCEQ Comments, *supra* note 164.

170. *See* Petitioners' Reply To Respondent United States Environmental Protection Agency's Response In Opposition To Petitioners' Emergency Motion For A Stay Pending Review, No. 10-1092 (D.C. Cir. Jan. 7, 2011).

171. Hodel v. Va. Surface Mining & Reclamation Ass'n, 452 U.S. 264 (1981) (upholding a scheme of conditional preemption); FERC v. Mississippi, 456 U.S. 742 (1982) (upholding a scheme of conditional preemption).

172. *See supra* text accompanying note 85.

173. *Hodel*, 452 U.S. at 265; *FERC*, 456 U.S. at 778.

were handed down in the 1980s before *Garcia*, which overruled *National League of Cities* and thereby undermined the doctrinal foundation of both cases.[174] Then, in due course, *Garcia* was all but overruled by *New York* and *Printz*, both of which nonetheless reaffirmed *Hodel* and *FERC*.[175] So what have all these decisions left settled, exactly?

Let's take a closer look. In *Hodel*, the Supreme Court upheld a new federal law that sought to regulate surface coal mining.[176] The law established environmental performance standards for surface coal mining operations, and provided for each state to establish a regulatory program in accordance with the standards and guidance provided in the law, subject to federal approval,[177] as with a SIP under the CAA.[178] For states that failed to secure approval of the state program, the law required the federal agency (Department of the Interior) to establish a federal regulatory program to implement the law,[179] as with a FIP under the CAA.[180]

The case was on appeal to the Supreme Court directly from the federal district court[181] which had struck down the law. The district court had relied heavily on the federalism doctrine of *National League of Cities*:

> In *National League of Cities* the Court held that the minimum wage and overtime pay provisions of the Fair Labor Standards Act could not be applied to states and subdivisions of states. The fact that wages and hours of state employees were commerce or affected commerce was not questioned. The

174. Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528 (1985). In *Garcia*, the Court said, "The controversy in the present cases has focused on the third *Hodel* requirement—that the challenged federal statute trench on 'traditional governmental functions.' The District Court voiced a common concern: 'Despite the abundance of adjectives, identifying which particular state functions are immune remains difficult.' Just how troublesome the task has been is revealed by the results reached in other federal cases . . . . We find it difficult, if not impossible, to identify an organizing principle that places each of the cases in the first group on one side of a line and each of the cases in the second group on the other side." *Id.* (internal citations omitted).

175. 505 U.S. at 167; 521 U.S. at 926.

176. 452 U.S. 264 (1981) (upholding the Surface Mining Control and Reclamation Act of 1977, 30 U.S.C. §§ 1201 *et seq*).

177. 30 U.S.C. § 1253 (1977).

178. 42 U.S.C. §7410(a) (1977).

179. 30 U.S.C. §1254 (1977).

180. 42 U.S.C. § 7410(c) (1977).

181. At the time, 28 U.S.C. § 1252 (1977) provided for direct appeals to the Supreme Court from any federal court decision striking down a federal law in any suit to which the U.S. was a party. Section 1252 was repealed in 1988.

> Tenth Amendment, however, was found to preclude the exercise of commerce power "in a fashion that would impair the States' 'ability to function effectively in a federal system.'" In reaching that conclusion, the Court carefully distinguished the authority of Congress under the commerce clause to regulate "wholly private activity," from the authority to enact regulations that were "directed, not to private citizens, but to the States as States." According to the Court, the exercise of commerce power in the former is limited only by the requirement that "'the means chosen by (Congress) must be reasonably adapted to the end permitted by the Constitution,'" whereas its exercise in the latter is proscribed by the Tenth Amendment when matters "essential to (the States') separate and independent existence" are involved. Applying the holding of National League of Cities, congressional action alleged to be in contravention of the Tenth Amendment must be scrutinized to determine whether the legislation is directed to the states as states and usurps an "integral governmental function."[182]

The district court arguably articulated the proper Tenth Amendment test, namely whether "legislation is directed to the states as states and usurps an 'integral governmental function.'"[183] But it was on less firm ground when it set out to apply the standard. "The issue before the court, therefore, is whether the surface mining act is directed to the states as a sovereign entity, displacing its role as a decision-maker in areas of traditional governmental services, or whether the act is directed to private activity."[184] The former, according to the court, was prohibited, but the latter was not.[185]

A test focused on whether the federal law impinged on an area of "traditional governmental services" clearly was not the essential test of *National League of Cities*, which focused instead on the more essential question of whether the federal law "force[d] directly upon the States [Congress's] choices as to how essential decisions regarding the conduct of integral

182. Va. Surface Mining & Reclamation Ass'n v. Andrus, 483 F. Supp. 425, 432 (W.D. Va. 1980) (internal citations omitted).

183. *Id.*

184. *Id.*

185. *Id.* Outside the context of "traditional governmental services," the court noted, "[t]he induced-coerced distinction [of *Steward Machine Co. v. Davis*] can still be used to determine the validity of legislation affecting nontraditional governmental functions." *Id.* at 432 n.6. This confusing twist was thankfully relegated to a footnote.

government functions are to be made."[186] The district court's analysis was on untenable ground from that point forward. In striking down the Surface Mining Act, the court cited a variety of ways it "[displaced the state's] role as a decision-maker in areas of traditional governmental services."[187] "The Commonwealth [of Virginia] is deprived of its right to dictate whether this land could be better used for some other purpose."[188] "[T]he reclamation provisions adversely affect land values."[189] The required remedial landscaping "is economically infeasible and physically impossible." "[T]he state has been deprived of its choice as to how best to protect Virginia's environment."[190]

But the problem with the Surface Coal Mining Act was not that it displaced state authority in areas of traditional governmental functions, which is entirely permissible on preemption grounds, but rather that it *subverted* the instrumentalities of state government to accomplish federal ends. That is how the law threatened the "States' ability to function effectively in a federal system" and their "separate and independent existence."[191]

The district court had distilled from *National League of Cities* a promising articulation of the protections offered to state governments by the Tenth Amendment. But it so misapplied that standard to the facts that the subsequent reversal by the Supreme Court was almost inevitable.

Reversing the district court in *Hodel*, the Supreme Court observed that the Act simply did not require states to do anything.[192] According to the majority, "[t]he most that can be said is that the Surface Mining Act establishes a program of cooperative federalism that allows the States, within limits established by federal minimum standards, to enact and administer their own regulatory programs, structured to meet their own particular needs."[193] "Congress could constitutionally have enacted a statute prohibiting any state regulation of surface

186. Nat'l League of Cities v. Usery, 426 U.S. 833, 855 (1976).

187. Va. Surface Mining & Reclamation Ass'n v. Andrus, 483 F. Supp. 425, 432 (W.D. Va. 1980)

188. *Id.* at 434.

189. *Id.*

190. *Id.*

191. *Nat'l League of Cities*, 426 U.S. at 851–52.

192. Hodel v. Va. Surface Mining & Reclamation Ass'n, 452 U.S. 264, 288 (1981).

193. *Id.* at 289.

coal mining. We fail to see why the Surface Mining Act should become constitutionally suspect simply because Congress chose to allow the States a regulatory role."[194]

But if *Hodel* was correct to reverse the district court, it also failed to apply the Tenth Amendment standard actually articulated in *National League of Cities.* Instead, it elaborated a highly permissive three-part test: In order to run afoul of the Tenth Amendment, the federal law must (1) "regulate[] the 'States as States'"; (2) "address matters that are indisputably 'attribute[s] of state sovereignty'"; and (3) "impair their ability 'to structure integral operations in areas of traditional governmental functions.'"[195] It upheld the Act because the first of the three requirements was not satisfied.[196]

*Hodel*'s gloss on *National League of Cities* had little to do with the latter's most essential holding, namely that the Tenth Amendment protects "States' ability to function effectively in a federal system" and their "separate and independent existence" within that system.[197]

*Hodel*'s failure to apply the structural federalism argument articulated in *National League of Cities* carried forward into *FERC v. Mississippi,*[198] in which the Court upheld the Public Utility Regulatory Policies Act of 1978 (PURPA).[199] PURPA required States to consider federal regulatory standards while the states regulated public utilities within their jurisdictions.[200] It required states to regulate pursuant to federal directives or face federal regulatory preemption.[201] *FERC* held that the requirement to consider the Act's guidance did not run afoul of *National League of Cities* because "no state authority or nonregulated utility is required to adopt or implement the specified rate design or regulatory standards."[202] The Court noted precedents that "in effect directed state decisionmakers to take or refrain from taking certain actions."[203]

194. *Id.* at 290.
195. *Id.* at 287–88.
196. *Id.* at 288.
197. Nat'l League of Cities v. Usery, 426 U.S. 833, 851–52 (1976).
198. FERC v. Mississippi, 456 U.S. 742 (1982).
199. Pub L. No. 95-617, 92 Stat. 3117 (1978).
200. *Id.*
201. *Id.*
202. *FERC,* 456 U.S. at 749–50.
203. *Id.* at 762.

By focusing on the simple question of whether state implementation of federal requirements remained a matter of state prerogative, *Hodel* and *FERC* made the same mistake as *Dole.* In all three cases, the Court failed to apply the Tenth Amendment standard of *National League of Cities* as a principle of protection for the "States' ability to function effectively in a federal system" or their "separate and independent existence."[204]

*New York* and *Printz,* the cases that arguably abandoned *Garcia,* each distinguished the congressional legislation in both *Hodel* and *FERC* as falling outside the categorical prohibition on federal commandeering, because states retained a theoretical choice to refuse obeisance to federal preferences.[205] But neither *Hodel* nor *FERC* asked the questions later raised in *New York* and *Printz,* namely whether the congressional legislation left state legislators able to respond to local preferences, whether it kept the federal government accountable for federal policies, and whether the congressional scheme could be squared with the "structural framework of dual sovereignty" that created a need to protect "States' ability to function effectively in a federal system" and their "separate and independent existence" in the first place.[206]

The tacit resurrection of *National League of Cities* implied in the tacit overruling of *Garcia* treated *Hodel* and *FERC* as proper applications of the former, apparently because when *Garcia* overruled *National League of Cities,* it took *Hodel* and *FERC* down too.[207] But in fact, the approach taken in *Hodel* and *FERC* is far closer to that of *Garcia* than to that of *National League of Cities. Hodel* and *FERC* abandoned the most important aspect of the Tenth Amendment standard articulated in *National League of Cities,* and thereby marked decisive steps on the road to *Garcia.* A proper application of the structural federalism doctrine at the heart of *New York* and *Printz* would find that both the "traditional governmental function" standard applied by *Hodel* and *FERC,* and the process federalism standard applied by *Garcia,* stood on equally untenable reasoning and that all three cases should be overruled.

204. Nat'l League of Cities v. Usery, 426 U.S. 833, 851–52 (1976).

205. New York v. United States, 505 U.S. 144, 167 (1992); Printz v. United States, 521 U.S. 898, 926 (1997).

206. *Nat'l League of Cities,* 426 U.S. at 851–52.

207. *See supra* note 175 and accompanying text.

A pure Tenth Amendment standard such as that articulated in *National League of Cities, New York,* and *Printz,* which asks whether federal legislation can be squared with the structural framework of dual sovereignty, would cast considerable doubt on the EPA's greenhouse gas SIP Call and similar cases of conditional preemption. Conditional preemption deprives state governments of the ability to represent the preferences of their constituents, not in areas of traditional governmental functions, but in actually crafting state policy. Where their "encouraged" (and invariably coerced) deference to federal preferences reduces the accountability of the federal government with respect to the costs and consequences of implementing its own policies, the "States' ability to function effectively in a federal system"[208] is clearly impaired. The national political process offers only this protection to federalism: It directly constrains the choices the federal government can make, by forcing it to be accountable for the costs and consequences of its own policies. It is precisely this accountability that conditional preemption allows the federal government to escape—through a subversion of state governments. That is why conditional preemption is more than just a lesser-included part of wholesale preemption, and that is why conditional preemption should be carefully checked by the federal courts.

## VI. TOWARDS A SEPARATION OF FEDERAL AND STATE GOVERNMENTS

The common aspect of *Dole, New York,* and *Printz* is that they distinguished between encouragement and coercion by looking to the effects on self-governance, accountability, and the individual rights protected by the diffusion of power. But if you reverse their logic, and start by looking at those effects, many of the supposedly permissible "encouragements" of cooperative federalism are clearly coercive.

In the case of conditional federal funds, searching for the point at which "pressure turns into compulsion" is a fool's errand.[209] Given the federal taxing monopoly, the loss of even a small amount of revenue to state governments forces the states to raise revenue elsewhere. Whether it's a dollar or $10 billion,

208. *Nat'l League of Cities,* 426 U.S. at 843.

209. South Dakota v. Dole, 483 U.S. 203, 211 (1987) (quoting Steward Mach. Co. v. Davis, 301 U.S. 548, 590 (1937)).

the citizens of the state lose either money or self-governing autonomy. Forcing states to make that choice is coercion.

In the case of conditional preemption, the states are forced to choose between two alternatives both of which entail the loss of self-governance, and one of which (where the state chooses to comply with federal conditions) also entails the loss of accountable government. Forcing states to make *that* choice is coercion.

The Supreme Court's progressive turn away from unbridled national majority rule, and back to the "structural framework of dual sovereignty,"[210] has been predicated on the Court's insistence that states must remain "independent and autonomous within their proper sphere of authority."[211] This marks a promising trend. But the Court's future federalism decisions will struggle with the contradictions and dilemmas of past precedents until it concludes that the pressure which the federal government puts on the states by indirection and "encouragement" is every bit as offensive to the structural framework of dual sovereignty as compulsion directly applied.

In the case of conditional federal grants, every dollar the federal government taxes away from a state's residents, only to transfer it back to the state government on condition of obeisance to the will of the national majority is a surrender of state sovereignty not merely purchased, but coerced. In the case of conditional preemption, every condition attached to federal "permission" for state regulation in an area the federal government could regulate itself is a shifting of the burden of regulation, taxation, and accountability from the federal government to the states. In the first case, the federal government would not be able to achieve its objective except through coercion, because it has no power to compel state regulation. In the second, the federal government could achieve its objectives by bearing the burden of regulation and taxation—and being accountable for the results, itself. But in neither case should the federal government be able to escape the burdens of implementing its own policies through the stratagem of dressing coercion of state governments up as encouragement.

210. Printz v. United States, 521 U.S. 898, 928, 932 (1997).

211. *Id.* at 929, n.14.

In future federalism cases, where the federal government "encourages" a state to forsake the preferences of its citizens by holding the hammer of the federal tax-and-spend power, or that of federal preemption, over the state government, the Court will hopefully recall its opinion in *Printz*:

> By forcing state governments to absorb the financial burden of implementing a federal regulatory program, Members of Congress can take credit for "solving" problems without having to ask their constituents to pay for the solutions with higher federal taxes. And even when the States are not forced to absorb the costs of implementing a federal program, they are still put in the position of taking the blame for its burdensomeness and for its defects.[212]

Cooperation between federal and state governments has proven enormously beneficial to our society. Consider the response to a hurricane or a wildfire. In those situations, federal and state officials stand shoulder to shoulder, on an equal footing, pooling resources and reaching decisions by consensus. The federal structure of our Constitution is not threatened when the partnership between federal and state governments is based on equality. But all too often the unequal power of the federal government is brought to bear on state governments through encouragement in the form of tangible pressure that by definition undermines the states' ability to remain "independent and autonomous within their proper sphere of authority."

The Supreme Court is hopefully working its way towards a doctrine of federalism that will offer real protection for the dual sovereignty of the states. But until such a doctrine comes to fruition, the federal essence of our Constitution has woefully little to protect it against the "arbitrary power" of unbridled national majority rule.

212. *Id.* at 930.

# PROTECTING SPEECH FROM THE HEART: HOW *CITIZENS UNITED* STRIKES DOWN POLITICAL SPEECH RESTRICTIONS ON CHURCHES AND CHARITIES

PAUL WEITZEL*



* J.D. U.C.L.A. 2009; paulweitzel@gmail.com. I thank my wife, Wendy Weitzel, for her constant support. I also thank David M. Carr, Michael W. Mosman, Joseph Mueller, Don Hansen, Adam Winkler, Zhimin Lin, John Dalton, Ajit Thind, Bryce Carr, Debbie Palmer, Simone Sema, Chris Truxler, Chris Jones, Richard Mo, Stuart Harris, Joe Wood, Eric Boorstin, Sam Brooks, and all my colleagues at Davis Polk & Wardwell. All errors are my own.

## I. INTRODUCTION

This article will argue that *Citizens United v. Federal Election Commission*[1] prohibits restricting the political speech of 501(c)(3) charitable organizations, in effect abrogating *Regan v. Taxation with Representation of Washington.*[2]

"The First Amendment 'has its fullest and most urgent application' to speech uttered during a campaign for political office."[3] The First Amendment grants individuals the right to endorse or denounce a candidate for office. After *Citizens United,* when individuals organize together as a for-profit corporation, they retain this right. But when individuals organize for a cause that is heartfelt, rather than economic, this right is eliminated. Ironically, this gag only restricts political speech, "speech that is central to the meaning and purpose of the First Amendment."[4]

In *Citizens United,* the United States Supreme Court recognized corporate speech rights. *Citizens United*'s plain language and reasoning also support allowing charitable organizations to engage in political speech. The traditional argument—that the free speech rights of these groups are not infringed because they can speak through affiliate organizations—was expressly rejected by *Citizens United.*[5] Likewise, the argument that these groups have sold their speech for subsidies is misplaced after *Citizens United*—corporations are subsidized by limited liability, as are 501(c)(4) organizations,

1. 130 S. Ct. 876 (2010).

2. 461 U.S. 540 (1983).

3. *Citizens United,* 130 S. Ct. at 898 (quoting Eu v. San Francisco Cnty. Democratic Cent. Comm., 489 U.S. 214, 223 (1989) (internal quotations omitted)).

4. *Id.* at 892 (citing Morse v. Frederick, 551 U.S. 393, 403 (2007)).

5. The counter position is masterfully explained by Roger Colinvaux and Miriam Galston. *See* Roger Colinvaux, Citizens United *and the Political Speech of Charities* 5–9 (Dec. 17, 2010) (unpublished manuscript), *available at* http://ssrn.com/abstract=1726407; Miriam Galston, *When Statutory Regimes Collide: Will* Wisconsin Right to Life *and* Citizens United *Invalidate Federal Tax Regulation of Campaign Activity?* 24 (George Washington Univ. Legal Studies Research Paper No. 499, Mar. 2010), *available at* http://ssrn.com/abstract=1572511. These authors argue that because the restriction arises from the tax code it is not a burden on speech, so it will be subject to a lower standard of review than strict scrutiny. This argument is based on the premise that the tax code merely removes a subsidy to political speech, rather than creating a penalty for speech. I address the implications of these arguments in Part III.B. *infra.*

which are also tax-exempt, but neither is barred from political participation.[6]

This article will argue that *Citizens United* implicitly recognized the right of 501(c)(3) organizations to endorse political candidates and engage in substantial lobbying. Part I will introduce 501(c)(3) organizations and the two major cases on point: *Regan v. Taxation with Representation of Washington*, which held that limits to the political speech of 501(c)(3) organizations were constitutional,[7] and *Citizens United v. Federal Elections Commission*, which held that some limits to the political speech applied to corporations were unconstitutional.[8] Part II will apply the Court's holding in *Citizens United* to 501(c)(3) organizations. Part III will discuss the possible consequences of removing the restriction.

## II. INTRODUCTION TO 501(C)(3) ORGANIZATIONS AND PRIOR CHALLENGES

### A. *Introduction to 501(c)(3) Organizations*

A 501(c)(3) organization is a non-profit, tax-exempt and tax-deductible organization.[9] Not all non-profit organizations are tax-exempt,[10] and not all tax-exempt organizations are eligible to receive tax-deductible donations.[11] Tax-exemption means the organization is not required to pay federal income taxes.[12] Tax-deductibility allows those who contribute to the organization to deduct those contributions from their income tax. 501(c)(3) organizations are both tax-exempt and tax-deductible and are prohibited from engaging in certain types of political speech.[13]

6. 26 U.S.C. § 501(c) (Supp. IV 2006).

7. 461 U.S. at 546.

8. 130 S. Ct. at 882.

9. 26 U.S.C. § 501(c).

10. "The word *non-profit* should not be confused with the term *not-for-profit* (although it often is). The former describes a type of organization; the latter describes a type of activity." Bruce R. Hopkins, THE LAW OF TAX-EXEMPT ORGANIZATIONS 4 n.4 (10th ed. 2011). *Compare* 26 U.S.C. § 501(c), *with* 26 U.S.C. § 183(a) (Supp. IV 2006).

11. *See* Hopkins, *supra* note 10, at 48–49.

12. 26 U.S.C. § 501(a), (c). Non-profit organizations are still subject to some form of tax. *See* Hopkins, *supra* note 10, at 6. For the scope of the tax exemption to 501(c)(3) organizations, *see* Parts II, III, and VI of Title 26, Subtitle A, Chapter 1, Subchapter F. *See also* 26 U.S.C. § 501(b).

13. 26 U.S.C. § 501(c).

A 501(c)(3) organization is a creation of the federal tax code.[14] It encompasses non-profit[15] groups "organized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, or educational purposes," as well as certain amateur sports organizations, and societies working to reduce child and animal abuse.[16]

The tax code imposes two broad speech restrictions on 501(c)(3) organizations. First, they may not have a "substantial part" of their activities dedicated to influencing legislation.[17] Second, they may not participate or intervene in political campaigns for or against a candidate for public office.[18] Along with these restrictions, 501(c)(3) organizations have two significant benefits: they are tax-exempt and contributions made to them are tax-deductible.[19]

### B. Regan v. Taxation with Representation of Washington

In *Regan*, the Supreme Court held that speech restrictions on 501(c)(3) organizations were constitutional.[20] Taxation with Representation of Washington (TWR) was a non-profit corporation organized to influence tax policy.[21] TWR was formed by merging a 501(c)(3) organization with a 501(c)(4) organization.[22] 501(c)(3) organizations are tax-exempt and tax-deductible, but cannot engage in substantial lobbying.[23] 501(c)(4) organizations are also tax-exempt, but are not tax-deductible and are not prohibited from engaging in substantial

14. *Id.* State law defines exemptions and deductibility from state taxes (e.g., income, property and sales). For a discussion of state tax exemptions and their limitations, *see, e.g.*, Camps Newfound/Owatonna, Inc. v. Town of Harrison, Me., 520 U.S. 564, 587 (1997) (labeling the use of tax exemptions in certain contexts as impermissible). This article will focus on the federal tax code because it is the source of the federal speech restriction.

15. The organization is non-profit if "no part of [its] net earnings . . . inure[] to the benefit of any private shareholder or individual." 26 U.S.C. § 501(c)(3).

16. *Id.*

17. *Id. But see id.* § 501(h) (specifying conditions under which tax-exempt status will be denied).

18. 26 U.S.C. § 501(c)(3). The contours of these restrictions have filled volumes. *See generally* Hopkins, *supra* note 10.

19. 26 U.S.C. § 170(a)(1), (c) (Supp. IV 2006); 26 U.S.C. § 501(c)(3).

20. Regan v. Taxation with Representation of Wash., 461 U.S. 540, 546 (1983).

21. *Id.* at 540.

22. *Id.*

23. 26 U.S.C. § 501(c)(3).

lobbying.[24] A 501(c)(4) organization is the corporate form usually used by lobbying organizations such as political action committees.[25]

Following the merger, the IRS denied 501(c)(3) status to TWR because it planned to engage in substantial lobbying.[26] The organization sued for a declaratory judgment, arguing that it qualified for 501(c)(3) status and that the restriction against substantial lobbying violated the First Amendment.[27] TWR also argued that denying deductibility to groups that engage in certain forms of speech "is in effect to penalize them for [such] speech."[28]

The unanimous Supreme Court made three key holdings. The first holding, which the concurrence referred to as the Court's "necessary assumption,"[29] was that TWR could reorganize into a 501(c)(4) organization, which is not tax-deductible but can still engage in substantial lobbying.[30] Because TWR had the option to reorganize as a 501(c)(4), the Court found that the organization as a whole was not entirely restricted from substantial lobbying. As a less drastic alternative to complete reorganization, the court said, "TWR can obtain tax-deductible contributions for its non-lobbying activity by returning to the dual structure it used in the past, with a § 501(c)(3) organization for non-lobbying activities and a § 501(c)(4) organization for lobbying."[31] In other words, a "501(c)(3) organization's right to speak is not infringed, because it is free to make known its views on legislation through its § 501(c)(4) affiliate without losing tax benefits for its non-lobbying

24. 26 U.S.C. § 501(c)(4). 501(c)(4) organizations face a practical limit to their lobbying: they must be "operated exclusively for the promotion of social welfare, or local associations of employees, the membership of which is limited to the employees of a designated person or persons in a particular municipality, and the net earnings of which are devoted exclusively to charitable, educational, or recreational purposes." *Id.* Excessive lobbying may disqualify an organization from claiming it is operated exclusively for these purposes.

25. *Id. See also* Fed. Election Comm'n v. Beaumont, 539 U.S. 123, 160 (2003) (stating section 501(c)(4) "covers some of the Nation's most politically powerful organizations, including the AARP, the National Rifle Association, and the Sierra Club").

26. *Regan*, 461 U.S. at 542.

27. *Id.*

28. *Id.* at 545 (quoting Speiser v. Randall, 357 U.S. 513, 518 (1958)).

29. *Id.* at 552 (Blackmun, J., concurring).

30. *Id.* at 544.

31. The Court previously stated, "Congress chose not to subsidize lobbying[, which is allowed only for a 501(c)(4) organization,] as extensively as it chose to subsidize other activities that non-profit organizations undertake to promote the public welfare." *Id.*

activities."[32] In effect, the option of creating a 501(c)(4) entity served as a safety valve to keep the speech restriction constitutional.

The Court's second key holding was that tax-exemption and tax-deductibility are both economically equivalent to cash subsidies; "A tax exemption has much the same effect as a cash grant to the organization of the amount of tax it would have to pay on its income. Deductible contributions are similar to cash grants of the amount of a portion of the individual's contributions."[33] The two tax benefits are distinguishable only by degree.[34]

The third holding addressed whether the tax was a penalty or whether it was merely withholding a subsidy. The Court found that taxation is not a penalty; exemption from taxation is a subsidy.[35] Because the government is not required to subsidize the exercise of a right, the government could withhold the subsidy when it was being used for political speech.[36] Tax exemptions and deductions are "a matter of grace [that] Congress can, of course, disallow . . . as it chooses."[37]

With these three holdings, the Court held that by placing speech restrictions on 501(c)(3) organizations, the government was merely declining to further subsidize charitable organizations to the extent they engaged in substantial lobbying. The "necessary assumption" was that lobbying activities could be channeled through a 501(c)(4) entity.

### C. Citizens United v. Federal Election Commission

*Citizens United* effectively abrogates *Regan.* Citizens United was a non-profit corporation.[38] Using donations from individuals

32. *Id.* at 553 (Blackmun, J., concurring).

33. *Id.* at 544. However, this assumes the donor will itemize deductions. Donations by low-income individuals may not be subsidized at all, which skews the benefit of deductibility toward organizations funded by wealthier donors.

34. *Id.* ("The system Congress has enacted provides this kind of subsidy to non-profit civic welfare organizations generally, and *an additional subsidy* to those charitable organizations that do not engage in substantial lobbying."(emphasis added)). *See also id.* ("Both tax exemptions and tax deductibility are a form of subsidy that is administered through the tax system.").

35. *Id.* at 549.

36. *Id.*

37. *Id.* at 549 (quoting Comm'r v. Sullivan, 356 U.S. 27, 28 (1958)) (internal quotations omitted) (alteration and omission in original).

38. Citizens United v. Fed. Election Comm'n, 130 S. Ct. 876, 886–87 (2010).

and corporations, it sought to distribute a film through cable television titled "Hillary: The Movie," a 90-minute documentary critical of a presidential candidate, during the primary elections.[39] Concerned that the release would violate election law, Citizens United sought a declaratory judgment against the Federal Elections Commission.[40] The district court found that the release would violate 2 U.S.C. § 441b, which prohibits corporations and labor unions from making "any broadcast, cable, or satellite communication" that "refers to a clearly identified candidate for Federal office" and is made within thirty days of a primary or sixty days of a general election.[41]

The case was appealed directly to the Supreme Court following a subsequent hearing by a three-judge panel.[42] The 5-4 majority reversed the district court, holding that section 441b violated the First Amendment's guarantee to free speech.[43] The Court found that "First Amendment protection extends to corporations,"[44] and "political speech does not lose First Amendment protection 'simply because its source is a corporation.'"[45] The majority applied strict scrutiny[46] to hold section 441b unconstitutional.[47]

One key argument rejected by the majority was that corporate speech was not actually limited because the corporation was still free to speak through a political action committee, or "PAC."[48] A PAC is a distinct organization that is specifically exempted from section 441b's corporate speech ban.[49] The Court held that being able to speak through a PAC does not cure the unconstitutionality of section 441b's speech restriction on

39. *Id.* at 887.

40. *Id.* at 886, 888.

41. *Id.* at 887 (citing 2 U.S.C. § 434(f)(3)(A) (Supp. IV 2006)). Violating this statute was a felony with a maximum penalty of five years in prison. 2 U.S.C. § 437g(d)(1)(A) (Supp. IV 2006).

42. *See* 28 U.S.C. § 1253 (Supp. IV 2006) (permitting direct appeal to the United States Supreme Court from "an order granting or denying . . . an interlocutory or permanent injunction in any civil action, suit or proceeding required by an Act of Congress to be heard and determined by a district court of three judges").

43. *Citizens United*, 130 S. Ct. at 917.

44. *Id.* at 899.

45. *Id.* at 900 (quoting First Nat'l Bank of Bos. v. Bellotti, 435 U.S. 765, 784 (1978)).

46. *Id.* at 898.

47. *Id.* at 917.

48. *Id.* at 897.

49. 2 U.S.C. § 441b(b)(2) (Supp. IV 2006).

corporations.[50] The Court noted that "PACs are burdensome alternatives; they are expensive to administer and subject to extensive regulations."[51] The Court then described the administrative requirements of a PAC, including staffing, recordkeeping, file retention, and monthly reporting to the FEC.[52] In addition, the Court noted that a corporation may not be able to set up a PAC in time to speak about a current campaign.[53] The Court held that "[s]ection 441b's prohibition on corporate independent expenditures is thus a ban on speech . . . . Were the Court to uphold these restrictions, the Government could repress speech by silencing certain voices at any of the various points in the speech process."[54] The Court struck down section 441b for impermissibly restricting speech.[55]

## III. *CITIZENS UNITED* PROHIBITS THE CURRENT SPEECH RESTRICTIONS ON 501(C)(3) ORGANIZATIONS

### A. Citizens United *Applies to 501(c)(3) Organizations*

In *Citizens United*, the Court dealt with corporations. Most 501(c)(3) organizations are organized as corporations, so applying *Citizens United*'s holding to non-profit and tax-deductible organizations is a natural fit. In fact, the plaintiff in *Citizens United* was a non-profit corporation.

The *Citizens Untied* opinion suggests the Court was aware that its holding would impact the non-profit sector. As an example of the perverse effects of the law banning corporate speech, the Court pointed out that the Sierra Club could not run an advertisement against a logging candidate, the NRA could not publish a book advocating the defeat of a candidate who endorsed a handgun ban, and the ACLU could not create a website supporting a candidate for defending free speech.[56] The Court emphasized that "[t]hese prohibitions are classic

50. *Citizens United*, 130 S. Ct. at 897–98.
51. *Id.* at 897.
52. *Id.*
53. *Id.* at 898.
54. *Id.*
55. *Id.* at 913.
56. *Id.* at 897.

examples of censorship."[57] Each of these organizations is affiliated with a 501(c)(3) organization.[58]

The holding in *Citizens United* makes no distinction between tax-deductible organizations and other entities using the corporate form. The Court affirmed that:

> The people determine through their votes the destiny of the nation. It is therefore important—vitally important—that all channels of communication be open to them during every election, that no point of view be restrained or barred, and that the people have access to the views of every group in the community.[59]

The Court did not qualify this statement—no group was to be restrained.

Speaking more directly, the Court held "[i]n the realm of protected speech, the legislature is constitutionally disqualified from dictating the subjects about which persons may speak and the speakers who may address a public issue."[60] Congress's ban on speech by 501(c)(3) organizations is Congress's attempt to dictate who may address a public issue. In light of *Citizens United*, that restriction is unconstitutional.

### B. Citizens United *Rejected the Reasoning in* Regan

*Citizens United* abrogates *Regan* by rejecting its reasoning in two ways. First, *Citizens United* rejects *Regan*'s alternate channel assumption—that an entity's right to speak can be exercised by speaking through an affiliate.[61] Second, *Citizens United* rejects the idea that a person can be forced to choose between some special advantages and the exercise of its fundamental rights, which

---

57. *Id.*

58. The Sierra Club is affiliated with the Sierra Club Foundation. *2004 Annual Report*, SIERRA CLUB, 37, http://www.sierraclub.org/foundation/downloads/2004annualreport.pdf (last visited Nov. 16, 2011). The NRA is affiliated with the NRA Foundation. *About the NRA Foundation*, NRA FOUND., http://www.nrafoundation.org/about/ (last visited Nov. 16, 2011). The ACLU is affiliated with the ACLU Foundation. *Donating to the American Civil Liberties Union and ACLU Foundation: What is the Difference?*, AM. CIVIL LIBERTIES UNION, http://www.aclu.org/donating-american-civil-liberties-union-and-aclu-foundation-what-difference (last visited Nov. 16, 2011).

59. *Citizens United*, 130 S. Ct. at 901 (quoting U.S. v. Auto. Workers, 352 U.S. 567, 593 (1957) (Douglas, J., dissenting)).

60. *Id.* at 902 (quoting First Nat'l Bank of Bos. v. Bellotti, 435 U.S. 765, 784–85 (1978)).

61. *Id.* at 897.

weakens *Regan*'s distinction between a penalty and a withdrawn subsidy.[62]

The reasoning in *Citizens United* is incompatible with the approach taken by *Regan*. If *Regan* survives *Citizens United*, their combined reasoning would allow the government to eliminate speech by any individual or group, thus allowing it to do indirectly what it cannot do directly.

### 1. *Citizens United* Rejected *Regan*'s Necessary Assumption as Unconstitutional

In *Regan*, the Court's opinion was premised on the "necessary assumption" that a 501(c)(3) could still speak by creating a 501(c)(4) affiliate.[63] It reasoned that as long as the organization was still able to speak through an affiliate, the ban was not a restriction on speech, but a permissible way to segregate funds. This alternate channel argument was rejected in *Citizens United*.[64]

Perhaps thinking of *Regan*, the Court in *Citizens United* directly addressed the argument that the restriction on speech was not a ban because corporations could still speak by creating affiliated PACs, and that the restriction merely prevented the corporation from commingling its political speech monies funded by donations expressly for that purpose with its general treasury.[65]

The Court rejected this reasoning:

> A PAC is a separate association from the corporation. So the PAC exemption from § 441b's expenditure ban, § 441b(b)(2), does not allow corporations to speak. Even if a PAC could somehow allow a corporation to speak—and it does not—the option to form PACs does not alleviate the First Amendment problems with § 441b.[66]

The Court thus rejected the alternate channel approach that it relied on in *Regan*. A 501(c)(4) organization is a separate association from the 501(c)(3). Its existence does not allow the 501(c)(3) to speak and does not alleviate the First Amendment

62. *Id.* at 905.

63. Regan v. Taxation with Representation of Wash., 461 U.S. 540, 544, 552 (1983); *see supra* Part II.C.

64. *Citizens United*, 130 S. Ct. at 897.

65. *Id.*

66. *Id.* The Court went on to explain that "[p]rohibited, too, are restrictions distinguishing among different speakers, allowing speech by some but not others . . . . Speech restrictions based on the identity of the speaker are all too often simply a means to control content." *Id.* at 898–99.

problems with the ban. The necessary assumption in *Regan*—that speech through a reorganized affiliate is sufficient—was struck down.

### 2. Persons Cannot Be Forced to Choose Between a Special Advantage and a Fundamental Right

*Regan*'s "necessary assumption" is not the only part of *Regan*'s reasoning stricken by *Citizens United*. *Regan* held that the ban on speech was acceptable because the funds in the 501(c)(3) organization were subsidized by tax-deductibility.[67] In *Regan*, the Court distinguished penalties from subsidies.[68] Congress was allowed to condition subsidies, but not penalties, on the waiving of the organization's fundamental rights.[69]

*Citizens United* rejected this holding by dropping the strained distinction between subsidy and penalty and finding plainly that a corporation cannot be forced to choose between a subsidy and a fundamental right:

> "[S]tate law grants corporations special advantages—such as limited liability, perpetual life, and favorable treatment of the accumulation and distribution of assets." This does not suffice, however, to allow laws prohibiting speech. "It is rudimentary that the State cannot exact as the price of those special advantages the forfeiture of First Amendment rights."[70]

Yet that is exactly what is currently required of 501(c)(3) organizations. They have been granted a special advantage, tax-deductibility, for which the state seeks to exact a "forfeiture of its First Amendment rights."[71] The Court's holding could not be more clear.

### 3. Applying *Regan*'s Reasoning After *Citizens United* Would Allow Congress to Prevent Any Speech

If *Regan* survives *Citizens United*, then Congress could silence corporations by creating two forms of for-profit corporations: those that are able to speak, and those that are not. Congress could raise corporate taxes to the point of shareholder protest,

67. *Regan*, 461 U.S. at 544.

68. *Id.*

69. *Id.*

70. *Citizens United*, 130 S. Ct. at 905 (quoting Austin v. Mich. Chamber of Commerce, 494 U.S. 652, 658–59, 680 (1990)) (internal citations omitted).

71. *Id.*

say 95%, but offer a subsidized rate to those that waive their First Amendment rights. In this way, corporations could be silenced through a de facto ban on corporate speech. Congress could justify its actions by saying it simply did not want to subsidize corporate political speech. The penalty/subsidy argument accepted by *Regan* would allow Congress to do indirectly what it could not do directly.

In *Citizens United*, the Court analyzed the constitutionality of the speech ban by considering whether the law would be constitutional if applied to individuals. "If § 441b applied to individuals, no one would believe that it is merely a time, place, or manner restriction on speech. Its purpose and effect are to silence entities whose voices the Government deems to be suspect."[72] Further, it stated that "[t]he Court has thus rejected the argument that political speech of corporations *or other associations* should be treated differently under the First Amendment simply because such associations are not 'natural persons.'"[73]

A similar comparison demonstrates the absurdity of *Regan* surviving *Citizens United*. Under the *Regan* logic, the government could vary tax rates depending on a person's individual political activity. There could be a tax for those who exercise their right to speak and a different tax for those who do not. Congress could justify this as merely choosing not to subsidize political speech by individuals. But, in effect, a citizen would be required to pay additional taxes to exercise a fundamental right. This would be a shadowy reincarnation of the poll taxes stricken in the 1960s.[74]

### C. Citizens United *Also Rejected* Regan*'s Policy Concerns*

Under the policy rationales stated in *Citizens United*, it makes sense to allow 501(c)(3) organizations to speak. First, these organizations have in-the-trenches experience on many important policies. Second, they do not carry the same risks as speech by for-profit corporations. And, third, removing the

72. *Id.* at 898.

73. *Id.* at 900 (emphasis added) (internal citations omitted).

74. *See* Harper v. Va. Bd. of Elections, 383 U.S. 663, 670 (1966) (holding that poll taxes are unconstitutional, because "where fundamental rights and liberties are asserted . . . classifications which might invade or restrain them must be closely scrutinized and carefully confined").

speech restrictions would eliminate the additional cost and chill already put on these organizations.

### 1. Expertise and Experience of 501(c)(3)s

By banning speech by charities, we lose the voices of some of the most passionate and knowledgeable groups in our society.[75] We exclude from welfare debates those running the soup kitchens. We exclude from environmental debates those most dedicated to conservation. We exclude from foreign policy debates those who work to heal our soldiers at war and who pray for peace. The price is too steep. "The Government has 'muffle[d] the voices that best represent the most significant segments of the economy.'"[76] Just as banning corporate speech muffles some of the most knowledgeable voices about our economy, banning charities from speaking muffles the voices of those who care passionately about helping others and those who care passionately about political issues. "The remedy of 'destroying the liberty' of some factions is 'worse than the disease.' Factions should be checked by permitting them all to speak and by entrusting the people to judge what is true and what is false."[77]

### 2. Granting Speech Rights to 501(c)(3)s Creates Fewer Risks Than Granting Speech Rights to Corporations

In *Citizens United*, the Court addressed the risks of allowing corporations to speak. These risks are no greater with charities, and often do not exist at all. Because charities are funded by donations, there is a far lower risk of forcing unwilling shareholders to sponsor speech they disagree with; far lower risk of aggregating massive amounts of wealth that is disproportionate to the contributor's belief in the message; and far lower risk of a loss of faith in democracy. There is little reason to expect that foreign money would be channeled through 501(c)(3) organizations any more than it is funneled

75. *Citizens United*, 130 S. Ct. at 912 ("On certain topics corporations may possess valuable expertise, leaving them the best equipped to point out errors or fallacies in speech of all sorts, including the speech of candidates and elected officials.").

76. *Id.* at 907 (alteration in original) (quoting McConnell v. Fed. Election Comm'n, 540 U.S. 93, 257–58 (2003) (Scalia, J., dissenting in part)).

77. *Id.* (quoting THE FEDERALIST NO. 10, at 130 (James Madison) (Benjamin Fletcher Wright ed., 1961)).

through 501(c)(4) organizations.[78] And while corporations were only banned from speaking for 60 days before an election, the ban on 501(c)(3) organizations is year-round.[79]

### 3. Granting Speech Rights to 501(c)(3)s Alleviates the Current Chill

Allowing political speech by 501(c)(3) organizations would also prevent some of the circumvention and difficulties with the gray areas the Court was concerned about in *Citizens United.* 501(c)(3) organizations are allowed to discuss political issues, but cannot endorse any candidate. However, "the distinction between discussion of issues and candidates and advocacy of election or defeat of candidates may often dissolve in practical application."[80] Because it is so difficult to delineate between the two, 501(c)(3) organizations are likely to be chilled from discussing issues for fear they may lose their tax-deductible status.

This difficulty is compounded in the YouTube generation of politics where retaining a lawyer to prescreen an internet advocacy video beforehand could cost more than producing the video itself. As production becomes less expensive and more agile, the relative cost of compliance increases and disproportionately affects organizations subject to speech restrictions.

## IV. POSSIBLE CONSEQUENCES OF STRIKING THE SPEECH RESTRICTIONS APPROVED IN *REGAN*

If linking tax-deductibility to limited speech is now unconstitutional, there are two ways the statute could be corrected. The Court could either strike down the speech restrictions on 501(c)(3) organizations, or the Court could strike down both the speech restrictions and the tax-

78. There is no reason to believe that foreign contributors would have a greater incentive to donate to a 501(c)(3) than a domestic contributor would. On the contrary, a foreign contributor would likely have less need of tax-deductibility in the United States.

79. *Compare* 2 U.S.C. § 434(f)(3)(A)(i)(II)(aa) (Supp. IV 2006) (defining prohibited electioneering communications as any broadcast communication that is made within 60 days of a general election and refers to a clearly identified candidate for federal office), *with* 26 U.S.C. § 501(c)(3) (Supp. IV 2006) (exempting corporations, charities, and foundations from taxes provided that they never attempt to intervene in any political campaign or attempt to influence legislation).

80. *Citizens United,* 130 S. Ct. at 909 (citing Buckley v. Valeo, 424 U.S. 1, 42 (1976)).

deductibility of 501(c)(3) organizations, making them similar to 501(c)(4) organizations. Because charities are subsidized in order to promote charitable work,[81] rather than as hush money, the better-reasoned approach is to strike down only the speech restrictions.

### A. *Tax-Deductibility Should Remain In Place*

Governments have a long tradition of using tax breaks to encourage charitable activity.[82] Federal tax subsidies to charities predate most of the tax code.[83] Tax subsidies by states began before the American Revolution.[84]

In contrast, the restrictions prohibiting endorsement of candidates by 501(c)(3) organizations only emerged in 1954, when then-Senator Lyndon B. Johnson added an amendment to the Revenue Act of 1954 in order to limit the power of his opponent in the coming primary election, who was backed by a charitable organization.[85] The ban was proposed on the floor of the Senate, without the benefit of committee review.[86] If the purpose of the speech restriction was to favor one candidate over the other, the time-tested practice of tax-deductibility ought not be thrown out with it.[87]

Tax-deductibility for charities also serves several policy goals.[88] Subsidizing charity encourages generosity, which helps counter the free-rider problem inherent to the public goods charities

81. *See* G.S.G., Annotation, *Exemption of charitable organization from taxation or special assessment*, 34 A.L.R. 634 (1925) (explaining the legal rationales upon which the exemption is based).

82. *See generally* Walz v. Tax Comm'n of N.Y., 397 U.S. 664, 675–80 (1970) (describing the ways that colonial governments and the early U.S. Congress exempted churches from property tax).

83. Tariff Act of 1894, ch. 349, § 32, 28 Stat. 556 (1894).

84. *See Walz*, 397 U.S. at 677–78, 682 (1970) (noting that tax subsidies of churches were widespread in colonial days and in the early United States); Joseph V. Sliskovich, *Charitable Contributions or Gifts: A Contemporaneous Look Back to the Future*, 57 UMKC L. REV. 437, 446 n.37 (1989) (explaining that American charity and gift law originated in pre-American Revolution English law).

85. Hopkins, *supra* note 10, at 608.

86. Hopkins, *supra* note 10, at 608 n.6 (citing 100 Cong. Rec. 9604 (1954)).

87. See Colinvaux, *supra* note 5, at 5–9, for an extensive history of Senator Johnson's amendment.

88. *But see* Stanley Surrey, *Tax Incentives as a Device for Implementing Government Policy: A Comparison with Direct Government Expenditures*, 83 HARV. L. REV. 705, 707–11 (1970) (arguing that a tax deduction is generally inferior to a direct subsidy as a means of achieving policy goals, because tax deductions are less equitable and more difficult to develop administer).

provide.[89] Subsidizing charity may also change an organization's focus in a way that increases the amount of charitable work provided. For example, Jill Horwitz found that not-for-profit hospitals were more likely to offer unprofitable services than for-profit hospitals.[90] Eliminating tax breaks for charities because we are afraid of what they might say would set back all of these policies, subjecting them to the market failures the tax breaks were designed to mitigate.

### B. *Flooding the 501(c)(3) Form*

One concern with allowing 501(c)(3) organizations to speak while retaining their tax-deductibility is that many other groups may begin to organize themselves under section 501(c)(3) just for tax-deductibility. This is a danger, but it will be mitigated by the other restrictions on 501(c)(3) organizations. 501(c)(3)s must be:

> [O]rganized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, or educational purposes, or to foster national or international amateur sports competition (but only if no part of its activities involve the provision of athletic facilities or equipment), or for the prevention of cruelty to children or animals[.][91]

Because this substantially limits the range of organizations that could fit within this description, the use of the 501(c)(3) form will remain somewhat limited.[92]

In addition, Congress could pass new legislation to create a framework that increases tax advantages for charitable *activities*, rather than for charitable *organizations*. The law already makes

89. David M. Schizer, *Subsidizing Charitable Contributions: Incentives, Information and the Private Pursuit of Public Goals* 2 (Columbia Law & Econ., Working Paper No. 327, 2008), *available at* http://ssrn.com/abstract=1097644.

90. Jill R. Horwitz, *Why We Need the Independent Sector: The Behavior, Law, and Ethics of Not-For-Profit Hospitals*, 50 UCLA L. REV. 1345, 1367–68 (2003).

91. 26 U.S.C. § 501(c)(3) (Supp. IV 2006).

92. There is also a limitation that "no part of the net earnings" of the 501(c)(3) organization "inures to the benefit of any private shareholder or individual[.]" 26 U.S.C. § 501(c)(3). This will probably not pose any real barrier though, because for-profit organizations can create separate 501(c)(3) entities, donate to them for tax deductions, then speak through them.

analogous distinctions,[93] though admittedly the issue becomes more difficult when dealing with religious charities.[94]

### C. *501(c)(3) Organizations and Religion*

Because religious groups are often organized as 501(c)(3) organizations, a number of commentators have considered whether tax-deductibility violates the Establishment Clause.[95] Other commentators have asked whether these speech restrictions violate the Free Exercise Clause.[96] Because so much has already been said on these topics, I will only briefly discuss the most relevant case, *Walz v. Tax Commission.*[97] For further depth, I recommend the articles cited above.

In *Walz,* the plaintiff argued that exempting religious organizations from paying property tax violated the Establishment Clause.[98] The Supreme Court recognized the difficulty of reconciling the Establishment Clause with the Free Exercise Clause, "both of which are cast in absolute terms, and either of which, if expanded to a logical extreme, would tend to clash with the other."[99] Because of this difficulty, the court found that the "First Amendment . . . does not say that in every and all respects there shall be a separation of Church and State."[100] The Court then listed various subsidies that flow from the government to churches, including tax-exemptions such as those for federal income tax and property tax[101] and indirect subsidies such as paying bus fares for students attending

93. *Compare* 26 U.S.C.§ 501(c) (describing tax-advantaged organizations), *with* 26 U.S.C.§ 183(a) (2006) (describing tax-advantaged activities).

94. *See* Walz v. Tax Comm'n of N.Y., 397 U.S. 664, 674 (1970) (stating that "[t]o give emphasis to so variable an aspect of the work of religious bodies would introduce an element of governmental evaluation and standards as to the worth of particular social welfare programs, thus producing a kind of continuing day-to-day relationship which the policy of neutrality seeks to minimize").

95. *E.g.,* Jennifer M. Smith, *Morse Code, Da Vinci Code, Tax Code and . . . Churches: An Historical and Constitutional Analysis of Why Section 501(c)(3) Does Not Apply to Churches,* 23 J.L. & POL. 41, 81 (2007) (arguing that treating churches differently than other tax-exempt organizations probably does not violate the Establishment Clause).

96. *E.g.,* Stephanie A. Bruch, *Politicking from the Pulpit: An Analysis of the IRS's Current Section 501(c)(3) Enforcement Efforts and How It Is Costing America,* 53 ST. LOUIS U. L.J. 1253, 1265 (2009); Smith, *supra* note 84, at 81.

97. 397 U.S. at 664.

98. *Id.* at 667.

99. *Id.* at 668–69.

100. *Id.* at 669 (internal citations and quotations omitted).

101. *Id.* at 676.

parochial as well as public schools and providing textbooks for parochial schools.[102]

The Court held that the property tax-exemptions were a "reasonable and balanced" attempt to guard against religious intolerance.[103] The Court also found that the purpose of the exemption was "neither the advancement nor the inhibition of religion," finding it relevant that the exemption had not "singled out one particular church or religious group or even churches as such; rather, it has granted exemption to all houses of religious worship within a broad class of property owned by non-profit, quasi-public corporations which include hospitals, libraries, playgrounds, scientific, professional, historical, and patriotic groups."[104]

The reasoning in *Walz* suggests that lifting the speech restriction against 501(c)(3) organizations while retaining tax-deductibility would not violate the Establishment Clause. *Walz* explained that:

> Adherents of particular faiths and individual churches frequently take strong positions on public issues including, as this case reveals in the several briefs amici, vigorous advocacy of legal or constitutional positions. Of course, churches as much as secular bodies and private citizens have that right. No perfect or absolute separation is really possible; the very existence of the Religion Clauses is an involvement of sorts—one that seeks to mark boundaries to avoid excessive entanglement.[105]

The Court found no distinction between speech by individuals or religious groups before concluding that the property tax-exemption was constitutional. According to the *Regan* Court, tax-exemptions and tax-deductions are both "form[s] of subsidy that [are] administered through the tax system.[106] So even according to *Regan*, lifting the speech restriction against 501(c)(3) organizations while retaining tax-deductibility would change only the degree of the subsidy, which is within Congress's power.[107] There is also no reason to suspect that the

102. *Id.* at 670–72.
103. *Id.* at 673.
104. *Id.* at 672–73.
105. *Id.* at 670.
106. Regan v. Taxation with Representation of Wash., 461 U.S. 540, 544 (1983).
107. *Id.* at 549.

other concerns found in *Walz*, such as singling out religion from other non-profits, would be present.

Because the change is merely a change of degree within the province of Congress, and no additional concerns mentioned in *Walz* are present, lifting the speech restrictions against 501(c)(3) organizations is unlikely to violate the Establishment Clause.

## V. CONCLUSION

After *Citizens United*, for-profit corporations can speak because "the worth of speech 'does not depend upon the identity of its source.'"[108] Overturning *Regan* would merely recognize that speech from the heart is as valuable as speech from the wallet.

Because the restrictions on political speech by non-profit organizations were premised on assumptions the Court rejected in *Citizens United*, *Regan* is effectively abrogated and 501(c)(3) organizations now have the right to fully engage in political speech.

108. Citizens United v. Fed. Election Comm'n, 130 S. Ct. 876, 904 (2010) (quoting First Nat'l Bank of Bos. v. Bellotti, 435 U.S. 765, 777 (1978)).

# THE NECESSITY OF FEDERAL INTELLIGENCE SHARING WITH SUB-FEDERAL AGENCIES

JASON B. JONES*



* B.B.A. Business Management, Southern Methodist University, 2008; J.D., The University of Texas School of Law, expected 2012. United States Army 2000–04, served in Tikrit, Iraq in 2003 and 2004 with A Co. 104th Military Intelligence Battalion and worked in the 4th Infantry Division, 1st Brigade S2. Many thanks to Professor Robert Chesney for his guidance in the development of this Note. My gratitude also goes to the editorial staff of the Texas Review of Law & Politics for their great work, especially Tim George and Nicholas Morrell. I am extremely grateful for the encouragement and support of Dr. Tamara Jones.

## I. INTRODUCTION

The tragic events of September 11, 2001, could have been avoided. The intelligence failures that made possible the hijacking of four commercial aircraft by foreign al Qaeda members operating inside the United States were not borne by any single agency or any single individual. Instead, they were part of a systemic intelligence failure—a failure that the 9/11 Commission addressed in the *Final Report of the National Commission on Terrorist Attacks Upon the United States* (9/11 Commission Report or Report).[1]

The 9/11 Commission Report recommended there be "unity of effort" in the sharing of intelligence.[2] Unity of effort is the "coordination and cooperation among all" intelligence agencies working "toward a commonly recognized objective."[3] In response to the Report and the events of 9/11, several laws and executive orders were adopted to implement the recommendations.[4] One focus of the new legislation was to expand the sharing of intelligence amongst the federal agencies and between federal agencies and state and local agencies. Under the new legislation, sub-federal agencies have the potential to play a larger role in the intelligence community, but that role is entirely dependent on the mechanisms put in place to encourage information sharing and to address risks of information sharing with state and local agencies.

The purpose of this Note is to analyze the past and current law governing the dissemination of national security information between federal, state, and local authorities, and to propose reforms. As a consequence of this focus, less emphasis will be placed on general policy questions related to information sharing, even though they are equally pertinent to discussions of

1. THE 9/11 COMMISSION REPORT, FINAL REPORT OF THE NATIONAL COMMISSION ON TERRORIST ATTACKS UPON THE UNITED STATES (2004) [hereinafter 9/11 COMMISSION REPORT].

2. *Id.* at 416.

3. Scott Lawrence, *Joint C2 Through Unity of Command*, JOINT FORCES Q., Autumn/Winter 1994–95, at 107.

4. The legislation passed in response to 9/11 and the 9/11 Commission Report includes the USA PATRIOT Act, the Homeland Security Act of 2002, the Intelligence Reform and Terrorism Prevention Act of 2004, and the Reducing-Over Classification Act. Executive Orders 13,354, 13,356, 13,526, and 13,549 were also adopted.

the legal basis for information sharing.[5] This Note assumes generally that information sharing is necessary for effective intelligence and has abstained from a policy analysis of that position. However, there are serious concerns with information sharing, and Part II gives a brief, non-exhaustive overview of some of those concerns.

Part II addresses the policy considerations of including state and local officials in intelligence operations. Part III explores the current cooperative arrangements between federal and sub-federal agencies and examines the attributes of each arrangement. Part IV discusses the history of information sharing among various intelligence agencies before 9/11. Part V is concerned with the changes to information sharing post-9/11 and in response to the 9/11 Commission Report. Part VI explores three recommendations: providing greater oversight and incentives for information sharing, establishing a central figure with authority to set standards for all agencies relating to information sharing, and improvements to existing cooperative arrangements.

## II. POLICY CONSIDERATIONS OF INCLUDING STATE AND LOCAL OFFICIALS IN INTELLIGENCE OPERATIONS

The destruction caused by the terrorist attacks on 9/11 instilled a sense that America was vulnerable to attacks from within. In the last several decades, homegrown terrorism[6] and radicalization[7] have increased, leading to a corresponding focus by federal intelligence agencies on counterradicalization.[8] With the shift of federal agencies toward counterradicalization and away from the Cold War intelligence bureaucracy, state and local agencies have particular advantages that can increase the

5. *See* Nathan Alexander Sales, *Mending Walls: Information Sharing After the USA PATRIOT Act*, 88 TEX. L. REV. 1795 (2010) (studying the merits of information sharing among intelligence agencies).

6. *See* Samuel J. Rascoff, *The Law of Homegrown (Counter)terrorism*, 88 TEX. L. REV. 1715, 1716–18 (2010) (defining homegrown terror as "the phenomenon whereby individual groups carry out attacks . . . within their native or adopted country or society").

7. *See id.* at 1718 (defining radicalization as "the process by which individuals or groups are socialized into a thought world that condones, valorizes, and ultimately may require acts of violence . . .").

8. *See id.* at 1718–19 (discussing the rise of homegrown terror and radicalization leading to a rise in counterradicalization). Counterradicalization tries to determine the reasons individuals join organizations and then attempts to counter those reasons and provide incentives not to join.

overall effectiveness of intelligence operations in combating homegrown terrorism. These advantages cannot be utilized without information sharing between federal agencies and state and local agencies. Professor Samuel J. Rascoff has identified several comparative strengths and weaknesses possessed by local intelligence agents.[9]

The first strength is called "Epistemic Federalism."[10] Because agencies approach issues from their particular perspectives, local agencies are more adept at seeing local factors of terrorism than are federal agencies.[11] This is especially important in countering homegrown terror if the notion that terrorist networks are not highly structured organizations, but rather a "loosely knit network" linking informal groups,[12] is true.[13] If this "bottom up" perspective is true, within Epistemic Federalism local agencies have several distinct structural advantages over their federal counterparts. Local agencies have comparatively large staffs, drawn from local populations with similar cultural and linguistic diversity as their areas of operation. Local agencies also "have a broad mandate . . . rather than circumscribed authority merely to enforce the law."[14] The role of the FBI and other executive branch officials is to enforce the law; however, local police have a much broader mandate "to serve and protect," which affords them the ability to help local citizens with their problems.[15] This in turn allows local agencies to develop ties with their communities that may be more challenging to develop with a narrower mandate. Thus, Epistemic Federalism allows local agencies to see terror operations differently than federal agencies, which creates a broader understanding of the threat and increases the effectiveness of counterradicalization.

9. I defer heavily to Professor Rascoff in this section because of his extensive study of the topic and real-world practical experience as Director of Intelligence Analysis for the New York City Police Department.

10. Rascoff, *supra* note 6, at 1726.

11. *Id.*

12. *Id.* at 1727.

13. Professor Rascoff calls the disorganized terrorist network theory the "bottom up" perspective. *Id.* at 1728.

14. *Id.* at 1730.

15. *See* Steven M. Cox, *Policing into the 21st Century*, 13 POLICE STUDIES: INT'L REV. POLICE DEV. 168, 168 (1990) (now titled POLICING: AN INT'L J. OF POLICE STRATEGIES & MGMT.) (highlighting the roles of municipal police that extend beyond law enforcement).

The second advantage associated with utilizing state and local agencies is coproduction and counterradicalization.[16] Counterradicalization implies "an intelligence effort that seeks out knowledge about social facts taking place within discrete communities, including information about individuals believed to be helpful to the authorities in pursuing their counterradicalization agenda."[17] Acquiring this sort of specialized intelligence requires utilizing coproduction,[18] '"the process through which inputs used to produce a good or service are contributed by individuals' who are 'clients' of [the] public good."[19] For coproduction to be effective, local citizens must be actively involved and local police are well-positioned to utilize their relationships with local communities to harvest information.[20]

A third advantage is that informal mechanisms and incentives may cause local police to have greater attentiveness to issues of basic civil rights during intelligence missions.[21] This is fostered through the accountability of elections or the "relationship of local police officials with the communities they secure."[22] Another argument is that police play multiple roles in the communities in which they operate, which creates incentives to be less intrusive and objectionable in their counterintelligence role.[23] However, not all commentators agree that local police can be held accountable for their counterterrorism operations,[24] which would weaken the argument that accountability benefits local citizens by ensuring attentiveness to their basic rights.

While state and local agencies bring specific advantages to the table, they also bring several inherent weaknesses. Local officials

16. Rascoff, *supra* note 6, at 1731.

17. *Id.* at 1732.

18. *Id.*

19. *Id.* Professor Rascoff adopts the definition of "coproduction" developed by Elinor Ostrom in *Crossing the Great Divide: Coproduction, Synergy, and Development, in* STATE-SOCIETY SYNERGY 85, 86 (Peter Evans ed., 1997).

20. *See* Rascoff, *supra* note 6, at 1732–34.

21. *Id.* at 1720.

22. *Id.* at 1738.

23. *Id.* Professor Rascoff calls this the "balanced portfolio" rationale. *Id.*

24. *See generally* Matthew C. Waxman, *Police and National Security: American Local Law Enforcement and Counterterrorism After 9/11*, 3 J. NAT'L SEC. L. & POL'Y 377, 391–99 (2009) (discussing the possible lack of accountability of local police forces in intelligence operations because local populations may not see direct benefits of counterintelligence work while their resources are being utilized for its operation and because counterintelligence operations are often secret and local populations will not know or understand what their officers are doing).

may have advantages in collecting local intelligence; however, one weakness is that they lack the analytical capacity to fully utilize the information they collect.[25] Professor Rascoff identifies three factors to support this proposition. First, counterradicalization requires the ability to comprehend and organize unrelated data points, and local agencies lack the analytical resources to perform this task.[26] Second, there are not mechanisms in place to vet intelligence collected at the local level and local agencies are unable to assess the accuracy of the information themselves.[27] Finally, there is currently no structure in place to connect the intelligence gathered by disparate local agencies and combine that information with federally collected intelligence information.[28] Until these challenges are addressed, local agencies will lack the analytical capacity to fully utilize the information they collect.

The second weakness identified is the lack of formal governance mechanisms ensuring basic rights are respected during intelligence operations by state and local agencies.[29] Consent decrees by federal courts "no longer effectively cabin police authority, and the internal guidelines that were promulgated to give them effect have similarly been relaxed."[30] Furthermore, legislative checks are generally not as effective at the local level and there is poor judicial review of intelligence matters.[31] Because of these challenges, there is generally little formal governance during local intelligence work to ensure basic rights are being protected.

A third weakness is that utilizing local agencies adds immense challenges to information sharing. There are an estimated 730,000 state and local full-time enforcement officers and between 13,500 and 19,000 state and local police agencies.[32] Expanding the information-sharing network to include all these individuals magnifies privacy risks,[33] security risks,[34] and civil

25. Rascoff, *supra* note 6, at 1720.

26. *Id.* at 1735.

27. *Id.*

28. *Id.* at 1735–36.

29. *Id.* at 1721.

30. *Id.* at 1741. For discussion of the cases relied on for this proposition, see Rascoff, *supra* note 6, at 1741 n.117.

31. Rascoff, *supra* note 6, at 1741.

32. Waxman, *supra* note 24, at 386, 380.

33. *Id.* at 390.

34. *Id.* at 391.

liberty concerns—discussed further below. Finally, to create an effective information-sharing network, some degree of standardization must exist.[35] Standardization is difficult to achieve given the varying degrees of sophistication, resources, and capabilities amongst agencies and the federal government's inability to compel reform at the local level.[36]

One sub-element of the third weakness is the privacy and security concerns that arise with increased information sharing. Privacy concerns are raised because information collected in one location will be distributed to individuals in multiple locations, perhaps with no relation to the original source.[37] With regards to security concerns, the possibility of classified information or information that is not being disclosed for security reasons being leaked increases with the greater number of officials having access, especially with the political pressure local police agencies may face.[38] Privacy and security issues are not the only concerns with increased information sharing; there are also substantial civil liberty concerns.

Opponents of expansive information sharing frequently discuss the civil liberty concerns inherent in such a system. A central concern is that by simply having information sharing programs in place, combined with intrusive governmental programs like wiretapping, they will have a chilling effect on individuals and infringe upon their First Amendment rights.[39] Another civil liberty concern is focused on the acquisition phase of intelligence. During the acquisition phase, there can be an invasion of privacy rights because every piece of information that a target sends is examined, not just the relevant intelligence information.[40] While intelligence acquisition raises civil liberty issues, greater concerns are raised during the actual sharing of

35. *Id.*

36. *Id.* For further information on the federal government's inability to dictate local reform, see *Printz v. United States*, 521 U.S. 898, 900 (1997) ("[T]he Federal Government may not compel the States to enact or administer a federal regulatory program.").

37. Waxman, *supra* note 24, at 390.

38. *Id.* at 391.

39. *See* ACLU v. NSA, 438 F. Supp. 2d 754, 768 (E.D. Mich. 2006) (discussing the plaintiffs' claims that warrantless electronic surveillance impeded their professional activities by chilling their speech or the speech of individuals integral to their work).

40. *See* William Pollak, *Shu'ubiyya or Security? Preserving Civil Liberties by Limiting FISA Evidence to National Security Prosecutions*, 42 U. MICH. J.L. REFORM 221, 259–60 (2008) (discussing invasion of privacy concerns during the acquisition of intelligence information through electronic surveillance).

information. As will be discussed later in the Note,[41] in response to 9/11, the government has continuously removed barriers to information sharing and created policies and programs to encourage widespread information sharing. This compounds civil liberty concerns by dramatically increasing the number of people who have access to information collected.

In sum, state and local agencies have significant inherent advantages that enhance the effectiveness of intelligence operations relating to both homegrown terrorism and counterradicalization. These advantages cannot be recreated through federal agencies, so there must be a partnership between federal agencies and state and local agencies. However, for state and local agencies to be effectively utilized, information sharing must necessarily occur, which has the potential to create the problems discussed above. Thoughtful and adequate safety measures must be put in place to address the problems and challenges identified. These advantages and disadvantages have been borne out in practice through cooperative arrangements between federal and sub-federal agencies.

## III. EXISTING COOPERATIVE ARRANGEMENTS

Federal agencies recognize the advantages state and local agencies provide in counterterrorism intelligence, and beginning in the 1980s,[42] several agencies have established cooperative arrangements utilizing state and local officials. The FBI created Joint Terrorism Task Forces (JTTFs), the Department of Homeland Security (DHS) sponsors fusion centers, and the Office of the Director of National Intelligence and the National Counterterrorism Center supports the Interagency Threat Assessment and Coordination Group (ITACG). Incorporating state and local agencies through cooperative arrangements may be a positive step, but each arrangement is imperfect and could be improved.

Since being established in 1980, JTTFs have continued to grow in both size and manpower, with over 4,400 agents

41. See *supra* Part IV (discussing governmental responses to 9/11).

42. Press Release, Fed. Bureau of Investigation, The Early Years: Celebrating 30 Years and the Beginning of New York's Joint Terrorism Task Force (Nov. 29, 2010), http://www.fbi.gov/newyork/stories/the-early-years/the-early-years (describing the first joint terrorism task force, which was created in New York in 1980 and was composed of ten members of the FBI and ten members of the NYPD).

nationwide today from more than 600 state and local agencies.[43] The purpose of the JTTFs is to facilitate greater communication between federal agencies and state and local agencies and to leverage the manpower and knowledge of local police forces.[44] The local agents are attached to the FBI, given access to classified information, and discouraged from communicating with and utilizing their home agency.[45] By effectively federalizing the agents and discouraging communication with their local agency, the local agents lose some of their inherent advantages discussed above and information sharing effectively ceases with state and local agencies, further weakening the advantages. While JTTFs were a positive first step, they are not ideal arrangements for counterterrorism purposes.

Fusion centers are supported by the DHS and designed to improve sharing of terrorism information between federal, state, and local authorities.[46] Fusion centers seem to have an advantage over JTTFs because state and local officials play a more substantial role.[47] In practice, unfortunately, fusion centers, like JTTFs, still have significant shortcomings. First, fusion centers have strayed away from their initial focus on terrorist activity and have instead become a center for "all threats and all hazards."[48] Second, fusion centers have been utilized simply as conduits of information and not centers for analysis.[49] During the earlier discussion of the challenges inherent in utilizing state and local agencies in intelligence operations, one of the main criticisms levied was the inability of state and local agencies to analyze information collected.[50] This becomes an even greater concern in the context of fusion

43. PROTECTING AMERICA FROM TERRORIST ATTACK: OUR JOINT TERRORISM TASK FORCES, http://www.fbi.gov/about-us/investigate/terrorism/terrorism_jttfs (last visited Nov. 23, 2011).

44. *Id.*

45. Rascoff, *supra* note 6, at 1743.

46. Michael German & Jay Stanley, AM. CIVIL LIBERTIES UNION, *What's Wrong with Fusion Centers?* 6, 9 (2007), http://www.aclu.org/pdfs/privacy/fusioncenter_20071212.pdf (last visited Nov. 23, 2011).

47. *See* Rascoff, *supra* note 6, at 1745 (discussing how the initiation of fusion centers by state and local agencies provides a more significant state and local presence).

48. *Id.* (quoting Ryan Singel, *Feds Tout New Domestic Intelligence Centers*, WIRED (Mar. 20, 2008), http://www.wired.com/threatlevel/2008/03/feds-tout-new-d (internal citations omitted).

49. *See id.* (discussing fusion centers exchanging rather than analyzing information).

50. *See supra* text accompanying notes 24–27 (discussing state and local officials' inability to analyze information they collect).

centers that act solely in an exchange capacity because intelligence sharing presupposes that the information being shared has been analyzed.[51] If extensive information is shared without analysis, it can lead to a flooding of the market, making intelligence work more challenging because large quantities of information must be sifted through. Recently, the DHS announced plans for more central control of fusion centers and to improve information analysis;[52] however, time will tell what effects the improvements have.

A third cooperative arrangement, supported by the Office of the Director of National Intelligence and the National Counterterrorism Center, is the ITACG. The ITACG embeds local officials within intelligence headquarters in Washington, D.C.[53] This model exposes the local officials to federally developed intelligence products and federal officials to counterterrorism issues faced by sub-federal officials.[54] While ITACG presents a better platform for information sharing than other cooperative arrangements, it suffers by adopting a federal-centric approach to intelligence production and treats sub-federal officials as mere consumers of intelligence and not co-producers.[55] Despite these protestations, ITACG still provides substantial benefits to local officials because they have an opportunity to provide feedback to help tailor intelligence products to suit their needs.

## IV. HISTORY OF THE SHARING OF INFORMATION AMONG VARIOUS INTELLIGENCE AGENCIES BEFORE 9/11

With the advantages to information sharing discussed in Part II, why are agencies reluctant to share information both horizontally and vertically? To begin answering that question, we

51. *See* Rascoff, *supra* note 6, at 1745.

52. *See I&A Reconceived: Defining a Homeland Security Intelligence Role. Hearing Before the Subcomm. on Intelligence, Information Sharing and Terrorism Risk Assessment of the H. Comm. on Homeland Security*, 111th Cong. (2009) (statement of Bart R. Johnson, Acting Under Secretary for Intelligence and Analysis, Department of Homeland Security), *available at* http://www.dhs.gov/ynews/testimony/testimony_1253802171234.shtm (discussing the DHS initiative to help fusion centers to gather, assess, analyze and share locally generated and national information and intelligence).

53. Rascoff, *supra* note 6, at 1724.

54. *Id.* at 1746.

55. *Id.* at 1747 ("[ITACG] perpetuates the flawed habit of regarding subnational participants principally as consumers of federal intelligence products, rather than as representatives of agencies with the capacity to gather and analyze intelligence alongside federal counterparts.").

must first understand the history of information sharing. This section focuses on the evolution of information sharing from the eve of World War II through 9/11, when information ceased to flow as a result of "the wall." The creation of "the wall" was not the fault of any single agency or individual. Instead it was the combination of a number of social events, legislative choices, and legal decisions.[56] Part IV is focused on explaining those events to lay a foundation for understanding why intelligence agencies hoard information.

### A. *Putting the Brakes on State and Local Agency Intelligence Collection*

By September 11, 2001, intelligence collection was almost exclusively a federal endeavor, but this was not always the case. Prior to World War II, large metropolitan police forces began collecting intelligence related to national security.[57] The New York City Police Department (NYPD) established a fifty-person intelligence squad.[58] The FBI, under J. Edgar Hoover's leadership, became concerned that the publicity generated by the program would cause citizens to transmit information concerning sabotage to the police rather than the FBI.[59]

In response to the NYPD's squad, Hoover brought the situation to the attention of the Attorney General and strongly urged the President to "issue a statement or request addressed to all police officials in the United States [] asking them to turn over to the FBI any information obtained pertaining to espionage, counterespionage, sabotage, and neutrality regulations."[60] The Attorney General's office immediately drafted a document to President Roosevelt and he released a statement. The statement had two main components: the FBI was to "take charge of investigative work in matters relating to espionage, sabotage, and violations of the neutrality regulations," and all law enforcement officers were to promptly

56. *See* Richard Henry Seamon & William Dylan Gardner, *The Patriot Act and the Wall Between Foreign Intelligence and Law Enforcement*, 28 HARV. J.L. & PUB. POL'Y 319, 323 & n.13 (2005).

57. Rascoff, *supra* note 6, at 1715.

58. *See* 1 NAT'L COUNTERINTELLIGENCE CTR., A COUNTERINTELLIGENCE READER: AMERICAN REVOLUTION TO WORLD WAR II 171 (Frank J. Rafalko ed., 2004), *available at* http://www.fas.org/irp/ops/ci/docs/ci1/chap4.pdf (discussing how the NYPD established a "special sabotage squad of fifty detectives").

59. *Id.*

60. Memorandum from FBI Director Hoover to Attorney General Frank Murphy (Mar. 6, 1939), *in* A COUNTERINTELLIGENCE READER, *supra* note 58, at 171.

turn over to the FBI "any information obtained by them relating to espionage, counterespionage, sabotage, subversive activities and violations of the neutrality laws."[61]

This statement illustrates that President Roosevelt did not envision a situation where federal agencies cooperated with state and local agencies in intelligence-related matters. The President made clear that the FBI was to take charge in this area, and state and local agencies were expected to turn over any information to the FBI. What the President did not say in this context is important: President Roosevelt did not discuss state and local agencies having any role in FBI intelligence operations other than acting as a conduit for any information obtained. There was no mention of the FBI sharing any information with state and local agencies. President Roosevelt effectively created a barrier preventing the sharing of information with state and local agencies.

If President Roosevelt's statement was the only impediment to sharing intelligence with state and local agencies, crafting a solution would be simple. However, subsequent events discussed below led to "the wall" of separation that prevented horizontal information sharing between federal intelligence agencies. While the following discussion is not directly related to the vertical sharing of information between federal agencies and state and local agencies, the discussion is important to understanding the barriers that were erected generally related to information sharing.

### B. *FISA and "The Wall"*

Much like the construction of a barrier built by the likes of Qin Shi Huang or Hadrian,[62] the metaphorical wall was not built overnight or with a single brick. Instead, it was created by the compilation of years of legislation, rules promulgated by the Justice Department, and natural organizational incentives of the different agencies. The origins of "the wall" can be traced to the practice of warrantless wiretapping.

In *Olmstead v. United States*, the Supreme Court considered whether the Fourth Amendment applied to the government's

61. Statement of President Roosevelt (Sept. 6, 1939), *in* A COUNTERINTELLIGENCE READER, *supra* note 58, at 172.

62. Qin Shi Huang built the first version of the Great Wall of China, and Hadrian built Hadrian's Wall, which marked the northern limit of Roman territory in Britain.

wiretapping activities.[63] The wiretapping involved in the case was conducted by intercepts located in a basement and on a street.[64] The court relied heavily on the lack of trespass involved in determining that there was no search and seizure, and therefore, that the Fourth Amendment was not implicated.[65]

Congress began regulating warrantless wiretapping in 1934 with its prohibition on the interception and dissemination of the contents of wire and radio communications under section 605 of the Federal Communications Act.[66] In *Nardone v. United States,* the prohibition was interpreted to apply to federal agents and the Court held that evidence obtained through wiretaps was inadmissible.[67]

While the prohibition seemed expansive, beginning with Roosevelt, presidents authorized warrantless electronic surveillance for national security purposes.[68] Roosevelt stated:

> I am convinced that the Supreme Court never intended any dictum in [*Nardone*] to apply to grave matters involving the defense of the nation.
>
> . . . .
>
> You are, therefore, authorized and directed . . . to secure information by listening devices directed to the conversation or other communications of persons suspected of subversive activities against the Government of the United States . . . .[69]

The practice of warrantless electronic surveillance for national security purposes was supported by Attorney General Tom Clark, who "advised President Truman of the necessity of using wiretaps 'in cases vitally affecting the domestic security.'"[70] The practice continued in organized crime and domestic security cases through at least the Johnson Administration.[71]

---

63. 277 U.S. 438 (1928).

64. *Id.* at 457.

65. *Id.* at 464.

66. 47 U.S.C. § 605 (1934).

67. Nardone v. United States, 302 U.S. 379, 384 (1937); *see also* Nardone v. United States, 308 U.S. 338, 340 (1939) (extending the earlier decision to exclude evidence indirectly obtained as the result of a prohibited interception).

68. Seamon & Gardner, *supra* note 56, at 330.

69. *See* S. REP. NO. 94-755, Book III, at 279 (1976) (author's emphasis removed) (citing Franklin D. Roosevelt, Confidential Memorandum for the Attorney General (May 21, 1940)).

70. United States v. U.S. Dist. Court for E. Dist. of Mich., S. Div., 407 U.S. 297, 311 n.10 (1972) (quoting Brief of the United States at 16–18).

71. *See id.* ("The nature and extent of wiretapping apparently varied under different administrations and Attorneys General, but, except for the sharp curtailment under

The Department of Justice (DOJ) also interpreted the Federal Communications Act and the *Nardone* decision broadly. The DOJ interpreted the Act and *Nardone* as merely prohibiting the divulgence of the contents of any intercepted communications outside the Federal establishment[72]—a blow to information sharing with sub-federal actors, which was already hindered by Roosevelt's decision addressed above. Courts seemed to support the DOJ's rationale and subsequently upheld the power to conduct warrantless wiretapping for purposes of national security as an inherent power of the President.[73] But while courts upheld the Executive's power to conduct warrantless wiretapping in certain situations, the Court focused on ensuring citizens' Fourth Amendment rights were not violated in the process.

The physical trespass analysis developed in *Olmstead* continued as law until the Court reexamined the issue in *Katz v. United States*,[74] where the Court shifted the analysis to whether the wiretapping violated a "reasonable expectation of privacy."[75] While *Katz* "established that the Fourth Amendment's protections applied to people rather than places or tangible things," the Court "explicitly reserved the issue of whether safeguards other than prior authorization by a magistrate would satisfy the Fourth Amendment in a situation involving the national security."[76] *Katz* highlighted the need for legal guidelines governing the use of electronic surveillance.

---

Attorney General Ramsey Clark in the latter years of the Johnson administration, electronic surveillance has been used both against organized crime and in domestic security cases at least since the 1946 memorandum from Clark to Truman.").

72. *See* S. REP. NO. 95-604, at 10 (1977) ("[T]he Justice Department did not interpret the Federal Communications Act or the Nardone decision as prohibiting the interception of wire communications *per se*; rather only the interception and divulgence of their contents outside the Federal establishment was considered to be unlawful.").

73. *See* United States v. Brown, 484 F.2d 418, 426 (5th Cir. 1973) (discussing that the President may constitutionally authorize warrantless wiretaps for the purpose of gathering foreign intelligence because of the "President's constitutional duty to act for the United States in the field of foreign relations, and his inherent power to protect national security in the context of foreign affairs"); *see also* Zweibon v. Mitchell, 363 F. Supp. 936, 944 (D.D.C. 1973) *rev'd*, 516 F.2d 594 (D.C. Cir. 1975) ("It is within the constitutional power of the President acting through the Attorney General to gather intelligence by authorizing electronic surveillance relating to foreign affairs and deemed essential to protect this nation and its citizens against hostile acts of a foreign power.").

74. 389 U.S. 347 (1967).

75. *Id.* at 360 (Harlan, J., concurring). Justice Harlan's phrase "reasonable expectation of privacy" has been considered the best summarization of the Court's holding and has been widely cited as the test developed in *Katz*.

76. Viet D. Dinh & Wendy J. Keifer, *FISA and The Patriot Act: A Look Back and A Look Forward*, 35 GEO. L.J. ANN. REV. CRIM. PROC. iii, vii (internal quotations omitted).

In response to *Katz*, Congress passed Title III of the Omnibus Crime Control and Safe Streets Act (Title III).[77] Title III generally prohibits the government from conducting electronic surveillance except in limited situations after obtaining a court order.[78] Title III requires the government to get advance judicial approval for electronic surveillance to investigate crime; however, it expressly allows the continued use of electronic surveillance for purposes of national security.[79]

While Title III does not address whether and how a warrant should be obtained for intelligence investigations as distinct from criminal investigation, the Court in *United States v. United States District Court for the Eastern District of Michigan* (*Keith*) answered the question in regards to a domestic, non-criminal investigation.[80] The Court found that national security investigations of domestic entities could implicate First and Fourth Amendment rights.[81] Therefore, when intelligence-gathering operations involve domestic organizations, the government must have probable cause of criminal wrongdoing and seek a warrant—but not a conventional criminal warrant—before utilizing electronic surveillance.[82] The Court left the door open for Congress to pass legislation on the type of warrant required—which Congress never did. The Court specifically noted that it did not pass "judgment on the scope of the President's surveillance power with respect to the activities of foreign powers, within or without this country,"[83] leaving unanswered whether the government must obtain prior judicial

77. Pub. L. No. 90-351, 82 Stat. 197 (codified as amended at 18 U.S.C. §§ 2510–22 (2008)).

78. *Id.*

79. *See* Title III, § 802, 82 Stat. 214 (1968), *repealed by* Foreign Intelligence Surveillance Act of 1978, Pub. L. No. 95-511, § 201(c), 92 Stat. 1797 ("Nothing contained in this chapter . . . shall limit the constitutional power of the President to take such measures as he deems necessary to protect the Nation against actual or potential attack or other hostile acts of a foreign power, to obtain foreign intelligence information deemed essential to the security of the United States, or to protect national security information against foreign intelligence activities. Nor shall anything contained in this chapter be deemed to limit the constitutional power of the President to take such measures as he deems necessary to protect the United States against the overthrow of the Government by force or other unlawful means, or against any other clear and present danger to the structure or existence of the Government.").

80. 407 U.S. 297 (1972).

81. George P. Varghese, *A Sense of Purpose: The Role of Law Enforcement in Foreign Intelligence Surveillance*, 152 U. PA. L. REV. 385, 390 (2003).

82. *Id.* at 391.

83. 407 U.S. at 308.

approval in cases involving national security and foreign entities. Consequently, the language of *Keith* seemed to denote a tripartite system of scrutiny under the Fourth Amendment. The highest level scrutiny of the procedures and standards for electronic surveillance was reserved for "ordinary crimes," with a lower standard, still requiring judicial approval, for electronic surveillance for information related to domestic threats to national security, and the lowest level of scrutiny for foreign threats to national security.[84]

After *Keith,* a series of troubling events were exposed that led to the enactment of the Foreign Intelligence Surveillance Act (FISA) in 1978.[85] These events are pertinent in that they highlight the abuse of intelligence information and the effect these events had on shaping information sharing for the next several decades.

In January 1970, Christopher Pyle revealed that the U.S. Army was spying on the civilian population.[86] Through Pyle's revelations and subsequent events, it was disclosed that the Army engaged in surveillance during political rallies and maintained files on candidates for office.[87] This revelation became the basis of a lawsuit, *Laird v. Tatum,* that reached the Supreme Court in 1972 and garnered considerable news coverage.[88] Several months after *Laird* was decided, the Watergate scandal broke, implicating the Justice Department, FBI, CIA, and White House.[89]

Then in 1973, the CIA compiled a list of all CIA activities that could violate its Charter.[90] The documents, called the "Family Jewels," revealed a number of operations conducted within the United States, including operations to electronically monitor U.S. reporters and to gather intelligence on protest movements

84. Seamon & Gardner, *supra* note 56, at 332.

85. Dinh & Keifer, *supra* note 76, at ix.

86. Pyle, *CONUS Intelligence: The Army Watches Civilian Politics,* 1 WASHINGTON MONTHLY, Jan. 1970.

87. George C. Christie, *Government Surveillance and Individual Freedom: A Proposed Statutory Response to* Laird v. Tatum *and the Broader Problem of Government Surveillance of the Individual,* 47 N.Y.U. L. REV. 871, 872 (1972).

88. 408 U.S. 1 (1972).

89. *See generally* CARL BERNSTEIN & BOB WOODWARD, ALL THE PRESIDENT'S MEN (1974).

90. CIA, Family Jewels 00418 (1973), http://www.foia.cia.gov/ [hereinafter "Family Jewels"] (type "Family Jewels" in the Search Declassified Docs browser; then select "Family Jewels" in the results).

in the United States.[91] The CIA kept the Family Jewels classified in fear of the damage its release would cause;[92] however, the papers were eventually leaked and the CIA's fears realized. On December 22, 1974, the country awoke to a front-page article in the *New York Times* revealing information contained within the Family Jewels.[93]

The article stunned the country, and within the next three months, the Executive and both Houses of Congress established committees to address the charges. President Ford established the "Rockefeller Commission" on January 4, 1975 to investigate CIA activities in the United States.[94] The Senate established the Senate Select Committee to Study Government Operations with Respect to Intelligence Activities (Church Committee) on January 27, 1975.[95] The House created a House Select Intelligence Committee (Nedzi Committee) on February 19, 1975.[96]

The Church Committee was influential and reported a number of occasions where intelligence activities "exceeded the restraints on the exercise of governmental power which are imposed by our country's Constitution, laws, and traditions."[97] In response to these abuses, the Committee "recommended a strict and careful separation of domestic and foreign intelligence gathering."[98] The reports of these committees[99] and the events

91. *Id.* at 00021, 00182.

92. Daniel L. Pines, *The Central Intelligence Agency's "Family Jewels": Legal Then? Legal Now?*, 84 IND. L.J. 637, 642 (2009).

93. Seymour Hersh, *Huge C.I.A. Operation Reported in U.S. Against Anti-War Forces, Other Dissidents in Nixon Years*, N. Y. TIMES, Dec. 22, 1974 at 1.

94. U.S. President's Commission on CIA Activities Within the United States: Files, [1947-1974] 1975, http://history-matters.com/archive/contents/church/contents_church_reports_rockcomm.htm (last visited Dec. 9, 2011).

95. Church Committee Created, http://www.senate.gov (search "Church Committee Created," then select "1. US Senate: Art & History Home > Historical Minutes > 1964...") (last visited Sept. 15, 2011).

96. The Nedzi Committee was created in February 1975 and was replaced by the Pike Committee in July 1975. The Pike Committee's official report was never released and is still classified, although it was purportedly leaked to the press and published in its entirety. Pines, *supra* note 92, at 643.

97. S. REP. NO. 94-755, bk. II, at 2 (1976).

98. William C. Banks, *The Death of FISA*, 91 MINN. L. REV. 1209, 1226 (2007).

99. SENATE SELECT COMM. TO STUDY GOVERNMENTAL OPERATIONS WITH RESPECT TO INTELLIGENCE ACTIVITIES, FINAL REPORT, S. REP. NO. 94-755 (1976); REPORT TO THE PRESIDENT BY THE COMMISSION ON CIA ACTIVITIES WITHIN THE UNITED STATES 48 (1975) (Rockefeller Commission Report).

described above were a substantial factor in the passage of FISA.[100]

FISA addressed the question left unanswered in *Keith* and created a structure for collecting intelligence information related to foreign powers or agents thereof. In order to be approved under FISA, an application for a Foreign Intelligence Surveillance Court (FISC) order must be made by a Senate-confirmed Executive official working in the area of national security who certifies the "*purpose* of the surveillance is to obtain foreign intelligence information."[101] Courts interpreted this to mean obtaining foreign intelligence information must be the "primary purpose."[102] If the primary purpose test is satisfied, information can also be used in criminal prosecutions when certain conditions are met.[103]

Some commentators suggest that Congress intentionally designed FISA to ensure that information obtained by electronic surveillance would rarely be used in criminal proceedings.[104] By

100. *See, e.g.*, 124 Cong. Rec. 10,889 (1978) (statement of Sen. Bayh) ("[T]his bill is required . . . because of certain misconduct and abuse which are almost unbelievable.").

101. *See* 50 U.S.C. § 1804(a)(7)(B) (2000) (prior to the 2001 and 2004 amendments) (emphasis added).

102. *See* United States v. Truong Dinh Hung, 629 F.2d 908, 932 (4th Cir. 1980) ("[T]he executive should be excused from securing a warrant only when the surveillance is conducted 'primarily' for foreign intelligence reasons."); *see also* United States v. Duggan, 743 F.2d 59, 77 (2d Cir. 1984) ("The requirement that foreign intelligence information be the primary objective of the surveillance is plain not only from the language of § 1802(b) but also from the requirements in § 1804 as to what the application must contain."); United States v. Pelton, 835 F.2d 1067, 1076 (4th Cir. 1987) (discussing that the primary purpose of the surveillance was to gather foreign intelligence information); *accord* United States v. Johnson, 952 F.2d 565, 572 (1st Cir. 1991) (discussing that the investigation of criminal activity cannot be the primary purpose of the surveillance). For more detailed discussion on this issue, see Dinh & Keifer, *supra* note 76, at xi.

103. *Truong Dinh Hung*, 629 F.2d at 932. The state or federal government must give notice of its intended use to the person against whom the information will be used before it can be used in a criminal trial or any other proceeding. *See* 50 U.S.C. §§ 1806(c), 1806(d) (2000 & Supp. II 2002). This allows defendants to file motions to suppress the information. *See* 50 U.S.C. § 1806(e) (2000 & Supp. II 2002) ("Any person against whom evidence obtained or derived from an electronic surveillance to which he is an aggrieved person . . . may move to suppress the evidence obtained or derived from such electronic surveillance on the grounds that– (1) the information was unlawfully acquired; or (2) the surveillance was not made in conformity with an order of authorization or approval."); *see also* 50 U.S.C. § 1806(f) (2000 & Supp. II 2002) (discussing the use of *in camera* and *ex parte* review by a district court "to discover or obtain applications or orders or other materials relating to electronic surveillance or to discover, obtain, or suppress evidence or information obtained or derived from electronic surveillance under this chapter.").

104. *See* Seamon & Gardner, *supra* note 56, at 358 (suggesting the need for advance Attorney General approval at two stages implies Congress did not intend the procedure to be routinely used).

requiring advance approval from the Attorney General before both submission of the application for a FISA surveillance order and use of the information obtained under the order in a criminal proceeding, Congress intended this procedure to be used infrequently.[105]

Other commentators suggest that the language Congress chose clearly evinces an approval of using information obtained by electronic surveillance in criminal proceedings.[106] The argument generally proceeds on textual grounds by noting that the statute did not use the primary purpose language and explicitly discussed using the information in criminal proceedings.[107] Regardless of the opinion one espouses, the courts' restrictive interpretation of the purpose test led to a widely held view that the investigation of ordinary crime could not be the primary purpose of the surveillance[108] and instilled a distinction between ordinary criminal investigations and foreign intelligence investigations.

After the events between *Keith* and the passage of FISA, there was great concern about information sharing in general, but these concerns were exacerbated when dealing with state and local actors because of the concerns addressed in Part II.[109] What impact then, did Title III, *Keith*, and FISA have on the handoff of information between federal and sub-federal agencies?

The advanced judicial approval required by Title III in criminal investigations[110] creates a presumption against information sharing in the criminal investigation context. However, the standards with regards to information sharing related to national security are different. Information collected under FISA with its primary purpose related to foreign intelligence should be allowed to be freely shared with state and local agencies because it meets the requirements of FISA and does not implicate Fourth Amendment concerns. However, when the information relates to domestic entities, as is the case

105. *Id.*

106. Dinh & Keifer, *supra* note 76, at xii.

107. *See* 50 U.S.C. §§ 1806(b) ("No information acquired pursuant to this subchapter shall be disclosed for law enforcement purposes unless such disclosure is accompanied by a statement that such information . . . may only be used in a criminal proceeding with the advance authorization of the Attorney General.").

108. Dinh & Keifer, *supra* note 76, at xi.

109. *See supra* notes 32–38 and accompanying text (discussing inherent information sharing challenges with sub-federal agencies).

110. *Supra* notes 78–79.

with homegrown terror—where information sharing with sub-federal agencies provides the greatest benefits—there are more formative barriers. While courts are generally deferential to national security concerns,[111] the requirement of judicial approval in this context creates a presumption that information cannot flow unrestricted. This presumption was validated in the federal context—therefore, certainly applicable in the federal to sub-federal context as well—by the guidelines the DOJ enacted, which made information sharing between intelligence investigators and criminal investigators extremely challenging.

The DOJ plays a central role by supervising investigations using electronic information, including intelligence investigations not focused on prosecution. The information used by the Department and its method of collection was of primary importance to ensuring the principles of FISA were being followed. The DOJ eventually interpreted the "primary purpose" rulings as "saying that criminal prosecutors could be briefed on FISA information but could not direct or control its collection."[112] The rationale behind the separation was to guarantee the integrity of the FISA investigations by ensuring they remained primarily for intelligence purposes and not criminal prosecution purposes.[113] DOJ prosecutors understood they were not to improperly exploit the FISA process; however, the prosecution of Aldrich Ames for espionage in 1994 raised questions about whether the current policies were sufficient.[114] The DOJ was concerned that the judge would find that FISA warrants had been misused because of the numerous contacts between FBI agents and prosecutors.[115] Attorney General Janet Reno responded by implementing new policies to eliminate this concern.[116]

111. *See, e.g.*, United States v. Brown, 484 F.2d 418, 426 (5th Cir. 1973) (collecting cases where courts were deferential in national security matters).

112. 9/11 COMMISSION REPORT, *supra* note 1, at 78.

113. Memorandum from Janet Reno, U.S. Att'y Gen., to Director, Fed. Bureau of Investigation and U.S. Att'ys, 2, 6 (July 19, 1995) [hereinafter 1995 Guidelines], http://www.fas.org/irp/agency/doj/fisa/1995procs.html 2, 6 (July 19, 1995) [hereinafter 1995 Guidelines], http://www.fas.org/irp/agency/doj/fisa/1995procs.html ("The purpose of these procedures is to ensure that [foreign intelligence] and [foreign counterintelligence] investigations are conducted lawfully . . . .").

114. 9/11 COMMISSION REPORT, *supra* note 1, at 78.

115. *Id.*

116. *Id.*

If FISA laid the foundation for "the wall," DOJ erected "the wall" itself with the guidelines issued in 1995 (1995 Guidelines) regarding the conduct of investigations.[117] The 1995 Guidelines solidified the distinction between criminal investigations and foreign intelligence investigations, and made it substantially harder for collaboration between foreign intelligence and criminal investigators.[118] The Guidelines stated that "the FBI and Criminal Division should ensure that advice intended to preserve the option of a criminal prosecution does not inadvertently result in either the fact or the appearance of the Criminal Division's directing or controlling the FI [foreign intelligence] or FCI [foreign counterintelligence] investigation toward law enforcement objectives."[119] The 1995 Guidelines and the results they caused were eventually coined "the wall."[120]

While the 1995 Guidelines were only meant to control the sharing of information between the Criminal Division and the FBI, the effects were far-reaching. The 1995 Guidelines were not intended to stop information sharing, but in practice there was substantially less information sharing and coordination between the Criminal Division and the FBI.[121] As a result of the 1995 Guidelines and pressure from FBI leadership and the FISA court, barriers were built between FBI agents.[122] Compounding the situation, FBI Deputy Director Bryant cautioned agents that improper information sharing could be a "career stopper."[123] This combination of factors led FBI intelligence investigators in the field to believe they could not share FISA information with agents involved in criminal investigations at all, and eventually, that intelligence investigators could not share information of any kind with criminal investigators.[124] Consequentially, information sharing stopped. This played a decisive role in the intelligence failures that led to September 11, 2001.

A meeting that took place on June 11, 2001 provides a tragic example of the role "the wall" played in contributing to 9/11. The meeting was between an FBI agent (Jane), a CIA agent

117. 1995 Guidelines, *supra* note 113, at 6.
118. 9/11 COMMISSION REPORT, *supra* note 1, at 79.
119. 1995 Guidelines, *supra* note 113, at 6.
120. 9/11 COMMISSION REPORT, *supra* note 1, at 79.
121. *Id.*
122. *Id.*
123. *Id.*
124. *Id.*

(Dave), and several FBI agents who had been working on the USS Cole bombing.[125] Jane brought three photographs with her to the meeting that had been given to her by a CIA agent.[126] Jane had NSA signal intelligence information related to the photographs that she decided not to share with the other FBI agents because the NSA report contained caveats not to share the information with criminal investigators.[127] Unfortunately, those FBI agents had previously worked on the same case the NSA information related to and sharing this information would have made the agents very interested in learning more about a suspect named Mihdhar.[128] Dave also knew information about Mihdhar, but he did not share that information with anyone because he was not asked and because he believed that as a CIA analyst, he was not authorized to answer FBI questions.[129] Jane said she assumed that if Dave had any knowledge, he would have volunteered it.[130] As a result, no information about Mihdhar was shared at the meeting.[131] Mihdhar, it turns out, was the weak link in Al Qaeda's planning of 9/11 and his capture could have helped prevent the attack. Instead, Mihdhar flew into the United States two days after the meeting, but no one was looking for him.[132]

This example illustrates the importance of information sharing and the consequences that can result without it. This section has also illustrated that information can be abused by agencies, and the damage that results by improper collection. However, information properly collected—with sufficient policies in place to prevent its misuse—and appropriately shared can be valuable in creating actionable intelligence that saves lives, as could have been the case during the June 11 meeting discussed above. When agencies hoard information, whether because of agency problems or legal restraints, tragic consequences can result.

125. *Id.* at 268.
126. *Id.*
127. *Id.* at 269.
128. *Id.*
129. *Id.*
130. *Id.*
131. *Id.*
132. *Id.*

## V. THE AFTERMATH OF 9/11: REMOVING "THE WALL" AND INCENTIVIZING INFORMATION SHARING

Would better information sharing have prevented 9/11? That question is one that will be left to ponder for the ages. While it is unquestioned that information sharing could have helped alert agencies to particular dangers, that does not guarantee that the totality of the events could have been prevented. Regardless, 9/11 provided a deadly example of the crippling effects that a lack of information sharing can have on our intelligence agencies. In response, the 9/11 Commission was established and legislation was passed to improve information sharing.

### A. *The USA PATRIOT Act and the Homeland Security Act of 2002*

Within a week of the horrendous attacks that occurred on 9/11, the Bush Administration began drafting an early version of what became the USA PATRIOT Act (Patriot Act)[133] in order to tear down "the wall."[134] The final version of the Patriot Act made a significant change to FISA; section 218 changed the purpose requirement from the courts' interpretation of the primary purpose to a "significant purpose."[135] This change made it easier for law enforcement agents to obtain FISA warrants where the target was important for both intelligence and criminal prosecution purposes,[136] and helped lower "the wall."[137] Some members of Congress were concerned the amendment to FISA's purpose provision would allow government to circumvent the notice and probable cause requirements of the Fourth Amendment.[138] Despite the concerns, the Patriot Act passed quickly and with overwhelming support.[139]

133. USA PATRIOT Act of 2001, Pub. L. No. 107-56, 115 Stat. 272 (2002) (codified as amended in scattered titles of U.S.C.).

134. 9/11 COMMISSION REPORT, *supra* note 1, at 328.

135. *See* 50 U.S.C. § 1804(a)(6)(B) (2006) (effective Oct. 7, 2010) (stating that the new standard requires "that a significant purpose of the surveillance is to obtain foreign intelligence information").

136. Dinh & Keifer, *supra* note 76, at xv.

137. Seamon & Gardner, *supra* note 56, at 379.

138. *S. 1448, The Intelligence to Prevent Terrorism Act of 2001 and Other Legislative Proposals in the Wake of the September 11, 2001 Attacks: Hearing Before the S. Select Comm. on Intelligence,* 107th Cong. 30 (2001) (statement of Morton H. Halperin, Chairman, Center for National Security Studies and Council on Foreign Relations).

139. *See* Seamon & Gardner, *supra* note 56, at 377–79 (discussing the concerns of several members of the Senate and the passage in both houses with substantial support).

Another important provision in the Patriot Act, section 504, authorizes federal officers who conduct electronic surveillance to "consult with Federal law enforcement officers to coordinate efforts to investigate or protect against" attack or clandestine intelligence activities of a foreign power or its agent.[140] The Act went even further and ensured that coordination would not preclude certification of a significant purpose or the entry of a surveillance order.[141] These actions removed the foundation that created "the wall" and invited a reexamination of the 1995 DOJ Guidelines.

In keeping with the purpose of the Patriot Act, Attorney General Ashcroft implemented new guidelines (2002 Guidelines) to replace the 1995 Guidelines.[142] The policies require more interaction between law enforcement and intelligence agents and provide new procedures for FISA investigations conducted "primarily for a law enforcement purpose" but also having a "significant foreign intelligence purpose."[143]

The government then submitted the 2002 Procedures for *en banc* review by the FISC.[144] The FISC held against the government.[145] However, on appeal,[146] the Foreign Intelligence Surveillance Court of Review reversed the FISC decision[147] and concluded that FISA never contemplated a court inquiring into the government's purpose for seeking intelligence information

140. 50 U.S.C. § 1806(k) (2006).

141. *Id.*

142. Memorandum from John Ashcroft, U.S. Att'y Gen., to the Assistant Att'y Gen. of the Criminal Div., Dir. of the FBI, Counsel for Intelligence Policy & U.S. Att'ys (Mar. 6, 2002) [hereinafter the 2002 Procedures], http://www.fas.org/irp/agency/doj/fisa/ag030602.html.

143. *Id.* at I.

144. *In re* Sealed Case, 310 F.3d 717 (FISACt. Rev. 2002) (per curiam), *cert. denied*, 123 S. Ct. 1615 (2003).

145. *Id.*

146. The opinion was not appealed directly. Instead, the government brought an application for surveillance of a U.S. person under the 2002 Guidelines. Brief for the United States, *In re* the Foreign Intelligence Surveillance Court, 218 F. Supp. 2d 611 (FISA Ct. 2002) (No. 02-001), *available at* http://www.fas.org/irp/agency/doj/fisa/082102appeal.html. The FISC court imposed the same restrictions it had imposed against the 2002 Procedures during the *en banc* review. The government then appealed this ruling to the Foreign Intelligence Surveillance Court of Review (FISCR), its first ever appeal to FISCR. *In re Sealed Case*, 310 F.3d 717.

147. *In re Sealed Case*, 310 F.3d at 719–20.

at all.[148] In the end, the legality of the replacement of the primary purpose with the substantial purpose test in the Patriot Act was upheld.[149]

Another piece of legislation passed in response to 9/11 with several important information sharing provisions was the Homeland Security Act of 2002 (Homeland Security Act).[150] Section 892 of the Act states that "[u]nder procedures prescribed by the President, all appropriate agencies, including the intelligence community, shall, through information sharing systems, share homeland security information with Federal agencies and appropriate State and local personnel."[151]

This legislation marked a substantial change from the Patriot Act. The Patriot Act simply removed the legal barriers preventing intelligence sharing. However, the Homeland Security Act was an affirmative command that is significant in two regards. First, it commands that all agencies shall "share homeland security information . . . ."[152] Second, and more important for the purposes of this Note, the Act commands that the information be shared not just among the federal agencies, but also with "appropriate State and local personnel . . . ."[153] The Homeland Security Act signals a significant shift in Congress's approach to information sharing by acknowledging an agency behavioral problem: namely, that removing barriers alone may open the floodgates, but it does not ensure the water will flow.

Passage of the Homeland Security Act evinced Congress's concern with the information sharing problems that contributed to 9/11 and changes necessary to ensure information sharing. However, before continuing to pass legislation to address those failures, Congress established a commission to better understand how 9/11 happened and to avoid such a tragedy again.[154]

148. *Id.* at 723 ("It does not seem that FISA, at least as originally enacted, even contemplated that the FISA court would inquire into the government's purpose in seeking foreign intelligence information.").

149. *Id.* at 727.

150. Homeland Security Act of 2002, Pub. L. No. 107-296, 116 Stat. 2135 (codified as amended in scattered sections of 6 U.S.C.).

151. *Id.* § 892(b)(1).

152. *Id.*

153. *Id.*

154. *See* 9/11 COMMISSION REPORT, *supra* note 1, at XV (discussing the establishment of the 9/11 Commission, the report states that "[t]he nation was unprepared. How did this happen, and how can we avoid such tragedy again?").

### B. *The 9/11 Commission Report and Governmental Responses to the Commission's Recommendations: Executive Orders and the Intelligence Reform and Terrorism Prevention Act of 2004*

At the end of 2002, Congress and the President created the National Commission on Terrorist Attacks Upon the United States (9/11 Commission or Commission).[155] The purpose of the 9/11 Commission was to investigate "facts and circumstances relating to the terrorist attacks of September 11, 2001."[156] The findings of the Commission were released on July 22, 2004 in the 9/11 Commission Report. The Report made five major recommendations but only one recommendation is central to the focus of this Note: the unity of effort in the sharing of intelligence.

The Commission found that the biggest impediment to all-source analysis is the resistance to information sharing.[157] Before 9/11, in response to the events discussed previously, there was a pervasive belief among the agencies that a demonstrated 'need to know' be present before sharing. The Commission emphasized that the previous culture must be replaced with one in which the agencies believe they have a duty to disclose.[158] In order to accomplish this, the Commission made two minor recommendations.

First, information procedures should provide incentives for sharing, to restore a better balance between security and shared knowledge.[159] The Commission was primarily concerned with creating a system where reports were designed so that pertinent information could be quickly discovered and further information obtained if necessary. The Commission was only focused on the horizontal sharing of information across federal agencies.[160]

The Commission believed a decentralized network model—where agencies maintain their own databases but those databases are searchable across agency lines—would allow

155. Pub. L. No. 107-306 (2002).

156. *Id.*

157. 9/11 COMMISSION REPORT, *supra* note 1, at 416.

158. *Id.* at 417.

159. *Id.*

160. *See id.* at 418 ("We propose that information be shared horizontally, across new networks that transcend individual agencies.").

information to be shared horizontally.[161] However, the Commission did not make clear whether authority for developing the policies for sharing each agency's information should be made by a central figure, such as the proposed National Intelligence Director, or by the head of each agency.[162] An earlier complaint by the Commission that the then-current Director of Central Intelligence did not have the ability to set standards for the information infrastructure[163] lends support to the proposition that a central figure should develop policies for all agencies.

Second, the President should coordinate the resolution of legal, policy, and technical issues across agencies to create a "trusted information network."[164] Once again, the Commission was only focused on federal issues and did not discuss the inclusion of state and local agencies. However, as discussed below, the legislation and executive orders passed in response to the Report recognized the necessity of information sharing with state and local agencies.

The executive branch was first to respond to the 9/11 Commission Report. On August 27, 2004, just a month after the Report became public, President Bush issued two executive orders encouraging information sharing. Executive Order 13,354 laid out four policy goals, two of which directly pertain to information sharing among governmental agencies.[165] Those objectives include giving "the highest priority to ... the interchange of terrorism information among agencies, [and] the interchange of terrorism information between agencies and appropriate authorities of States and local governments."[166]

To achieve these objectives, the Order created a National Counterterrorism Center (NCTC), which was to serve as the primary analytical and planning center for the nation's counterterrorism activities.[167] This model assumes a hub-and-

161. *See id.* (discussing the current system and the decentralized network model).

162. *Id.* at 411.

163. *Id.* (discussing the three authorities critical for any agency head that the then-current DCI lacked).

164. *Id.* at 418. A trusted information network is a network where governmental actors can share vital information securely. Policies must also be in place to ensure civil liberty violations do not occur as a result of the network.

165. *See* Exec. Order No. 13,354 § 1(a)(ii)–(iii), 69 Fed. Reg. 53,589 (Sept. 1, 2004).

166. *Id.*

167. *Id.* § 3(a)–(b) (ordering that the National Counterterrorism Center will "(a) serve as the primary organization in the United States Government for analyzing and

spoke system where the NCTC acts as the hub and continuously receives and transmits information to all the spokes, including state and local agencies and precludes agencies sharing information directly as peers.[168] This Order was in direct conflict with the 9/11 Commission's decentralized network model recommendation.

Executive Order 13,354's envisioned structure was directly undermined by another executive order passed the same day: Executive Order 13,356.[169] While 13,356 espoused the same policy objectives regarding the interchange of terrorism information as 13,354,[170] 13,356 ordered the head of each agency to share terrorism-related information with the heads of the other agencies.[171] This method of achieving the policy objectives conflicted with 13,354's. 13,354 envisioned a hub-and-spoke system with the NCTC acting as a central clearing house for intelligence information, and 13,356 envisioned the heads of each agency sharing information with one another without going through a central clearing house.

Despite the contradiction with 13,354, Executive Order 13,356 has several other important features relating to information sharing. Following the understanding of the Homeland Security Act and the 9/11 Commission, 13,356 recognizes that agencies must be incentivized to share information. The Order directs agencies to create "appropriate arrangements providing incentives for . . . increased sharing of terrorism information . . . ."[172] The Order implemented another of the Commission's recommendations[173] by taking steps[174] to prevent barriers to the

integrating all intelligence possessed or acquired by the United States Government pertaining to terrorism and counterterrorism . . ." and "(b) conduct strategic operational planning for counterterrorism activities . . . .").

168. *See* Nathan Alexander Sales, *Share and Share Alike: Intelligence Agencies and Information Sharing*, 78 GEO. WASH. L. REV. 279, 301 (2010).

169. *See* Exec. Order No. 13,356, 69 Fed. Reg. 53,599 (Aug. 27, 2004). Executive Order 13,356 was later revoked, although its substantive provisions were mainly unchanged. For discussion, *see* Sales, *supra* note 168, at 301 n.140.

170. Exec. Order No. 13,356 § (1)(a)(ii)–(iii) ("[A]gencies shall . . . give the highest priority to . . . the interchange of terrorism information among agencies, [and] the interchange of terrorism information between agencies and appropriate authorities of States and local governments.").

171. *Id.* § 2.

172. *Id.* § 3(e).

173. 9/11 COMMISSION REPORT, *supra* note 1, at 417 (discussing the need to reduce over-classification of information that prevents information sharing).

174. Exec. Order No. 13,356 § 3(b)–(d) (directing agencies to produce multiple levels of information to allow varying degrees of access, "requiring terrorism information

distribution of information among agencies. The Orders were positive in that they encouraged information sharing among federal agencies, and with state and local agencies.[175] However, these positive aspects were diminished to some degree by the disagreement between the orders on the proper structure for information sharing.

Shortly after President Bush issued Executive Orders 13,354 and 13,356, Congress responded to the 9/11 Commission Report's recommendation with the passage of the Intelligence Reform and Terrorism Prevention Act of 2004 (IRTPA).[176] IRTPA is frequently discussed for its controversial step of creating a "Director of National Intelligence"[177] who oversees the intelligence community and serves as the chief intelligence adviser to the President.[178] However, the most important aspect of IRTPA for our purposes is section 1016, which established a new "information sharing environment" (ISE).[179]

Section 1016 directs the President to create an ISE, designate the structure to manage and operate the ISE, and determine and enforce guidelines related to the ISE.[180] The ISE envisioned by Congress under IRTPA confirms the information sharing initiatives under Executive Orders 13,354 and 13,356. IRTPA also establishes several new institutions related to the ISE.[181] However, IRTPA suffers from the same problem as some of the earlier legislation: it speaks in general platitudes without

---

to be shared free of originator controls," and "minimizing the applicability of information compartmentalization systems to terrorism information").

175. *See* Exec. Order No. 13,354 § 1(a)(ii)–(iii) (discussing the creation of the NCTC and state and local agencies as one of the constituents); *see also* Exec. Order No. 13,356 § 3 (mandating the preparation of terrorism information for maximum distribution within the intelligence community, in which state and local agencies were explicitly mentioned).

176. Intelligence Reform and Terrorism Prevention Act of 2004, Pub. L. No. 108-458, 118 Stat. 3638.

177. *See* Sales, *supra* note 168, at 299 (noting critiques of the creation of the Director of National Intelligence by Richard A. Posner. *See* RICHARD A. POSNER, PREVENTING SURPRISE ATTACKS: INTELLIGENCE REFORM IN THE WAKE OF 9/11, 51–56 (2005)).

178. Intelligence Reform and Terrorism Prevention Act of 2004 § 1011(a).

179. 6 U.S.C. § 485(a)(3) (2006). This section defines an ISE as "an approach that facilitates the sharing of terrorism and homeland security information, which may include any method determined necessary and appropriate."

180. *Id.* § 485(b)(1)(a)–(c).

181. *See* Sales, *supra* note 168, at 300 (discussing the creation of the ISE, Program Manager, and the Information Sharing Council).

providing substantive guidance on how the policies should be implemented and pursued.[182]

As directed by IRTPA, President Bush released the Guidelines and Requirements in Support of the Information Sharing Environment (ISE Guidelines).[183] The ISE Guidelines established five information-sharing guidelines,[184] two of which are important for purposes of this Note. Guideline 1 addresses the discrepancies between the models of information sharing envisioned in Executive Orders 13,354 and 13,356 and adopts the model described in Order 13,356.[185] Instead of a centralized intelligence database, the ISE guidelines adopted a decentralized approach to information sharing. Guideline 2 provided for the development of a common framework for information sharing, including with state and local agencies.[186] Once again, the importance of information sharing with sub-federal agencies was recognized.

Unlike other recent legislation addressed above, the ISE Guidelines actually provided substantive implementation procedures to encourage information sharing,[187] rather than simply touting the benefits of information sharing and appealing to agencies' interests.[188] They also hold senior managers and officials accountable if information sharing is not improved by including a performance evaluation element in the annual performance reviews.[189] The ISE Guidelines seek to address agency problems related to why intelligence agencies tend to hoard information rather than sharing it. All of the

182. *See id.* at 300–02 (discussing the lack of specificity in IRTPA and Executive Order 13,356).

183. Memorandum on Guidelines and Requirements in Support of the Information Sharing Environment: Memorandum for the Heads of Executive Departments and Agencies, 41 WEEKLY COMP. PRES. DOC. 1874 (Dec. 16, 2005) [hereinafter ISE Guidelines].

184. *Id.* § 2(a)–(e).

185. *See id.* § 2(a) (discussing the implementation of Executive Order 13,388, which revoked Executive Order 13,356 but left it substantively unchanged, see *supra* note 169 for more information).

186. ISE Guidelines *supra* note 82, at § 2(b) ("Recognizing that the war on terror must be a national effort, State, local, and tribal governments, law enforcement agencies, and the private sector must have the opportunity to participate as full partners in the ISE . . . .").

187. *See id.* § 3 ("Promoting a Culture of Information Sharing.").

188. *See* Sales, *supra* note 168, at 303 (addressing the difference between the ISE Guidelines and previous legislation).

189. *Id.* at 302–03.

changes discussed above significantly increased information sharing and the effectiveness of our intelligence agencies.

### C. *Legislation to Reduce Over-Classification Under the New Administration*

One of the concerns of the 9/11 Commission was the problem of over-classifying information.[190] Over-classification is a concern for information sharing because it limits the ability of federal agencies to share information about potential threats with sub-federal agencies—who may not have the proper security clearances. This issue has been addressed recently with a series of Executive Orders and the passage of the Reducing Over-Classification Act.[191]

Acknowledging the importance of the issue, President Obama issued Executive Order 13,526 within his first year of taking office.[192] The Order "prescribes a uniform system for classifying, safeguarding, and declassifying national security information . . . ."[193] The Order charged the head of each agency with establishing the distribution controls of classified information,[194] versus a central decision-making authority. This could potentially lead to information sharing problems if agencies create varying levels of control.

Executive Order 13,526 did not address the sharing of classified information with state and local agencies; however, Executive Order 13,549 was promulgated in August of 2010 to address this precise issue. The purpose of Order 13,549 is "to ensure that security standards governing access to and safeguarding of classified material are applied in accordance with Executive Order 13526" to information shared with state, local, tribal, and private sector entities.[195] To achieve its purpose, the Order applied the standards set forth in 13,526 to information sharing with sub-federal agencies. This Order shows a continued commitment by the Obama Administration to information sharing with state and local agencies. Similar to 13,526, Order 13,549 places the responsibility on the sponsoring

190. 9/11 COMMISSION REPORT, *supra* note 1, at 417–18.

191. Reducing Over-Classification Act, Pub. L. No. 111-258, 124 Stat. 2448 (2010).

192. Exec. Order No. 13,526, 75 Fed. Reg. 707 (Dec. 29, 2009).

193. *Id.*

194. *Id.* § 4.2(a).

195. Exec. Order No. 13,549, § 1.2, 75 Fed. Reg. 51,609 (Dec. 29, 2009).

agency for determining the eligibility of access for a state or local agency, as opposed to a central figure making such decisions.[196] Codifying Executive Order 13,526, Congress recently passed the Reducing Over-Classification Act.[197] In passing the Act, Congress made five findings, three of which relate to the negative effects that over-classification has on information sharing.[198] Another finding of fact acknowledges that the agencies authorized to make original classification decisions[199] "are responsible for developing, implementing, and administering policies, procedures, and programs that promote compliance with applicable laws, executive orders, and other authorities . . . ."[200] The Act provides another example that the agency-centric structure envisioned in Executive Order 13,356 has been adopted as the model henceforth. In sum, these measures should reduce the problems of over-classification and allow enhanced information flow.

## VI. RECOMMENDATIONS

Overall, the changes made in response to 9/11 and the 9/11 Commission Report have improved U.S. intelligence capabilities by improving information sharing. This has been accomplished through legislation removing barriers to information sharing and implementing policies and procedures that incentivize information sharing. While these changes have been substantial, improvements can still be made. This Note adopts three main recommendations: providing greater incentives and oversight for information sharing to overcome institutional design and agency problems that lead to information hoarding, adopting a centralized approach to the implementation of guidelines relating to the classification of information, and modifying existing cooperative arrangements to maximize their effectiveness. These changes will help encourage the positive

196. *See id.* at § 1.3(a) ("Eligibility for access to classified information by SLTPS personnel shall be determined by a sponsoring agency.").

197. Reducing Over-Classification Act, Pub. L. No. 111-258, 124 Stat. 2648 (2010).

198. *Id.* § 2(2)–(4).

199. *See* Exec. Order No. 13,526 § 1.3(a)(1)–(3) (noting that the authority to classify information originally may be exercised only by the President and Vice President, agency heads, and officials authorized to classify information by agency heads).

200. *See* Reducing Over-Classification Act § 2(5) ("Federal departments or agencies authorized to make original classification decisions or that perform derivative classification of information are responsible for developing, implementing, and administering policies, procedures, and programs . . . .").

steps taken since 9/11 and will further the recommendations developed in the 9/11 Commission Report.

The first recommendation is that incentives for information sharing must be expanded and continuously monitored to overcome agency self-interest to hoard information and encourage information sharing between federal and sub-federal agencies. Expanding incentives is not a novel idea; it was recommended in the 9/11 Commission Report.[201] It has also been the subject of discourse by many scholars. Congress and the Executive have made strides through recent legislation that explicitly provides incentives for information sharing and goes beyond simply removing barriers.[202] However, this is not enough. It is imperative that Congress continue monitoring the effectiveness of these incentives and make changes as necessary.

A study conducted on the development of three intelligence agencies noted that lack of oversight by Congress was a factor in agency behavioral problems.[203] Agencies are self-interested,[204] and Congress and the Executive must continue providing oversight and incentives to make it in the agencies' best interests to share information. The oversight and incentives must change the culture of information sharing from being a "career stopper" to being a "career strengthener." Information hoarding can result not only from agency behavioral problems, but also from dissimilar policies among varying agencies.

The second recommendation is to vest authority within a central figure to set information classification and sharing policies. As discussed in Part V, Section B, the 9/11 Commission recommended a decentralized network model where agencies maintain their own database and can search across agency lines; however, the Commission left unanswered whether authority for developing policies should be left with the heads of each agency or with a central figure.[205] An argument could be made that the

201. 9/11 COMMISSION REPORT, *supra* note 1, at 417.

202. Examples of recent legislation that do not simply remove barriers to information sharing, but instead provide affirmative commandments to share information include: the Homeland Security Act, Executive Order 13,356, and the ISE Guidelines. *See supra* Part V.

203. *See generally*, AMY B. ZEGART, FLAWED BY DESIGN: THE EVOLUTION OF THE CIA, JCS, AND NSC (1999) (examining the development of the Joint Chiefs of Staff, the Central Intelligence Agency, and the National Security Agency).

204. *See* Sales, *supra* note 168, at 281 (discussing the "iron law" of agency self-interest).

205. *See supra* text accompanying notes 160–62.

Commission preferred the central authority approach based on their complaint that the Director of Central Intelligence did not have the authority to set standards for the information infrastructure.

A central authority figure may be the approach the 9/11 Commission preferred; however, Congressional and Executive responses have been inconsistent in this area. Executive Order 13,356, promulgated by President Bush, ordered the DCI—a central authority figure—to create policies for information sharing within the intelligence community.[206] Similarly, in passage of IRTPA, Congress ordered the President to "issue guidelines for acquiring, accessing, sharing, and using information . . . ."[207] However, in both Executive Order 13,526[208] and the Reducing Over-Classification Act,[209] the authority to create guidelines is not vested within a central authority figure, but instead is delineated to the heads of each agency.

Certainly, an argument can be made that setting guidelines relating to the classification of information is integral to each agencies' operation and effectiveness. However, that argument does not overcome the importance of the free flow of information, and allowing agencies to establish different guidelines relating to classification can prevent information sharing and defeat the purpose of reducing over-classification in the first place. In response to the inconsistency in policy and the potential negative consequences, it is recommended that a central figure be given authority to establish guidelines—with consultation from agency heads—relating to the classification and sharing of information.

The final recommendation is to modify existing cooperative arrangements to maximize their effectiveness. While the current cooperative arrangements provide a solid foundation to build from, they each have inherent weaknesses that could be eliminated to make them more effective. There is not a one-size-fits-all approach that can be adopted because each agency has its

206. *See* Exec. Order No. 13,356, § 3 ("[T]he Director of Central Intelligence shall, in consultation with the Attorney General and the other heads of agencies within the Intelligence Community, set . . . common standards for the sharing of terrorism information by agencies within the Intelligence Community . . . .").

207. Intelligence Reform and Terrorism Prevention Act of 2004 § 1016(d)(1), Pub. L. No. 108-458, 118 Stat. 3638.

208. Exec. Order No. 13,526, 75 Fed. Reg. 730 (Dec. 29, 2009).

209. Reducing Over-Classification Act, Pub. L. No. 111-258, 124 Stat. 2448 (2010).

own unique objectives, but there are two central elements that would benefit all arrangements.

First, each arrangement would benefit by allowing state and local agents to continue coordinating with their home agency. This allows agents to stay abreast of any changes within their area of operation and any developments that their local partners have discovered. Removing that connection begins to erode the advantages that sub-federal agents bring to the cooperative arrangements.[210]

Second, state and local agents should be given a larger role in intelligence analysis. Through Epistemic Federalism, these agents bring unique perspectives and are more adept at seeing local factors of terrorism than federal agents.[211] Therefore, utilizing local agents in the analysis phase could enhance the effectiveness of the analysis and final product. Implicit in this recommendation is the need to train state and local agents in intelligence analysis. One of the weaknesses Professor Rascoff identified is that local agents lack the analytical capacity necessary to fully capitalize on information they collect.[212] If agents are given proper training and a larger role in intelligence analysis, the products produced and shared could become more effective.

## VII. CONCLUSION

Four decades ago, we were shown the devastating effects the abuse of information can have on our nation and our civil liberties. In response, changes were made to the ability to collect and share information. These changes led to the development of "the wall" and to a drastic decline in the sharing of intelligence information. Then, just over a decade ago, we were shown the devastating effects that can result when intelligence information is not shared. Once again, we have responded by making changes to our ability to collect and share information. These changes have removed many of the barriers preventing information sharing and have helped to overcome agencies' self-interest in hoarding information. But the job is not yet complete, and arguably never will be. We must remain vigilant

210. *See supra* Part II (discussing the advantages and disadvantages of state and local agents).

211. *See supra* notes 10–13 and accompanying text.

212. *See supra* notes 25–29 and accompanying text.

in overseeing agencies to ensure information is shared and monitoring the effectiveness of our current laws.


TEXAS REVIEW OF
LAW & POLITICS