PARTIAL UNCONSTITUTIONALITY

KEVIN C. WALSH*

Courts often hold legislation unconstitutional, but nearly always only part of the statute offends. The problem of partial unconstitutionality is therefore pervasive and persistent. Yet the exclusive doctrinal tool for dealing with this problem—severability doctrine—is deeply flawed. To make matters worse, severability doctrine is purportedly necessary for any workable system of judicial review. The accepted view is that severance saves: A court faced with a partially unconstitutional law must sever and excise the unconstitutional provisions or applications so that the constitutional remainder can be enforced going forward. Absent severance and excision, a law must fall in its entirety. This excision-based understanding of judicial review is supposedly traceable to Marbury v. Madison*. In fact, this attribution is anachronistic. Moreover, the prevailing view is wrong about the distinctive function of modern severability doctrine, which is not to save, but to destroy. This Article retrieves the original approach to partial unconstitutionality and develops a proposal for implementing a version of that approach. The proposal, displacement without inferred fallback law, is simultaneously ambitious and modest. It is ambitious because it proposes a shift in the general framework for judicial review in every case; it is modest because the proposed shift would change case outcomes in only a small set of highly consequential cases.*

 Assistant Professor, University of Richmond School of Law. Thank you to Todd Aagaard, Richard Fallon, Jim Gibson, Philip Hamburger, Benjamin Hatch, Scott Hemphill, Corinna Lain, James Bernard Murphy, Chaim Saiman, Carl Tobias, and various participants in the AALS FRC for comments on and criticisms of earlier versions of this Article. I would also like to thank Greg Bell, Carolyn Wagner, and Michael Matheson for indefatigable research assistance. I am grateful for financial assistance from the Villanova University School of Law and the University of Richmond School of Law.

INTRODUCTION

Many laws are unconstitutional, but few are entirely so. As a result, partial unconstitutionality is pervasive. Sometimes a statute's unconstitutionality resides in a discrete textual provision, such as a legislative veto in a complex immigration statute[1] or a private cause of action for gender-motivated violence in a massive crime bill.[2] More commonly, a statute's unconstitutionality inheres in some applications of the statute but not others, such as a speech restriction that is unconstitutional as applied to speech on public sidewalks but not elsewhere,[3] or heightened protection for religious liberty that is unconstitutional as applied to states and localities but constitutional as applied to the federal government.[4] Although often unnoticed, almost every instance of judicial review presents an issue of partial unconstitutionality in one form or another.

The problem of what to do with partially unconstitutional laws has a long pedigree. *Marbury v. Madison*, for instance, resulted in a holding of partial unconstitutionality.[5] The Court held unconstitutional just one part of the Judiciary Act of 1789 (Section 13) and just some applications of that part (Supreme Court issuance of writs of mandamus in the exercise of its original jurisdiction).[6] Chief Justice Marshall's reasoning in *Marbury* continues to frame the problem of partial unconstitutionality today. Following Marshall's reasoning, the question asked by a court engaged in judicial review is whether the act

---

[1] *See* INS v. Chadha, 462 U.S. 919, 956–59 (1983) (invalidating Immigration and Nationality Act, Pub. L. No. 414, § 244(b), (c), 66 Stat. 163, 216 (1952)).

[2] *See* United States v. Morrison, 529 U.S. 598, 627 (2000) (invalidating Violence Against Women Act of 1994, Pub. L. No. 103-322, § 40302(c), 108 Stat. 1796, 1941).

[3] *See* United States v. Grace, 461 U.S. 171, 183 (1983) (invalidating in part Act of August 18, 1949, Pub. L. No. 250, § 6, 63 Stat. 616, 617 as applied to public sidewalks surrounding Supreme Court building).

[4] *See* City of Boerne v. Flores, 521 U.S. 507, 534–36 (1997) (invalidating in part Religious Freedom Restoration Act (RFRA) of 1993, Pub. L. No. 103-141, § 6(a), 107 Stat. 1488, 1489 as applied to state and local governments).

[5] 5 U.S. (1 Cranch) 137, 176–80 (1803) (holding that portion of Judiciary Act of 1789 allowing Supreme Court "to issue writs of mandamus to public officers" conflicts with Constitution and is therefore void).

[6] Richard H. Fallon, Jr., *As-Applied and Facial Challenges and Third-Party Standing*, 113 HARV. L. REV. 1321, 1337 n.91 (2000) ("[T]he Court did not invalidate that linguistic unit of the [Judiciary Act of 1789] insofar as it authorized the Supreme Court to issue writs of mandamus in the exercise of its *appellate* jurisdiction.").

of the legislature is "void."[7] A holding of unconstitutionality is, therefore, a holding of voidness. If a legislative act is void, how can any portion of it be enforced after a judicial pronouncement of unconstitutionality?

Current law and scholarship answer that severability doctrine is the exclusive way to deal with partial unconstitutionality. Severability doctrine governs whether a court may first separate out or "sever" the unconstitutional provisions or applications of a law, and then subtract or "excise" them, so the constitutional remainder can be enforced going forward. Thus conceived, this judicial operation creates a new law that consists of the old law "minus" its unconstitutional provisions or applications.[8]

The lodestar for this severability determination is legislative intent. Because a court must not "use its remedial powers to circumvent the intent of the legislature,"[9] the court must ask before severing: "Would the legislature have preferred what is left of its statute to no statute at all?"[10] If the answer is "no," then the court should not sever but instead should declare the statute entirely invalid and enjoin its enforcement in all applications.

As the doctrinal formula for determining severability reveals, the required inquiry into legislative intent is unlike many other interpretive inquiries, in that it asks what the legislature would have done, not what the legislature actually did.[11] Thus, the doctrine leaves courts with significant discretion in deciding whether and how to sever and excise statutes. Legislatures that enact partially unconstitutional laws may not foresee their constitutional flaws and may not articulate a

---

7 *Marbury*, 5 U.S. (1 Cranch) at 177.

8 Michael C. Dorf, *Fallback Law*, 107 COLUM. L. REV. 303, 305 (2007) ("[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."); *see also* Richard H. Fallon, Jr., *The Linkage Between Justiciability and Remedies—And Their Connections to Substantive Rights*, 92 VA. L. REV. 633, 674–75 (2006) ("[T]he normal presumption of constitutional law is that statutes are 'severable' (or, synonymously, 'separable'): Even if a statute has unconstitutional components or would be unconstitutional as applied to particular facts, the unconstitutional elements or applications can be severed from the valid ones and the valid ones enforced.").

9 Ayotte v. Planned Parenthood of N. New England, 546 U.S. 320, 330 (2006) (quoting Califano v. Westcott, 443 U.S. 76, 94 (1979)).

10 *Id.* Additionally, the Court assumes that "Congress could not have intended a constitutionally flawed provision to be severed from the remainder of the statute if the balance of the legislation is incapable of functioning independently." Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 684 (1987).

11 *See* United States v. Booker, 543 U.S. 220, 246 (2005) (stating that severability depends on "what 'Congress would have intended' in light of the Court's constitutional holding") (citation omitted).

preference about how to cure any constitutional flaws that they do foresee.

Those who trust the judiciary to work things out by discerning the legislature's unstated intent about a matter that it never addressed when legislating can take heart from the current approach to partial unconstitutionality. The rest of us should be concerned. The problem is not simply that severability doctrine vests courts with substantial discretion, but that it does so in an area of the law in which the stakes are high.

The deployment of severability doctrine can be highly consequential because the doctrine allows a court to declare *entirely* invalid a law that is only *partially* unconstitutional.[12] The doctrine may also metastasize invalidity beyond unconstitutionality even when it yields the conclusion that particular statutory provisions are severable. For example, in the Supreme Court's recent high-profile deployment of the doctrine in *Booker v. United States*, the Court "severed and excised" two statutory provisions—one that made the Federal Sentencing Guidelines binding and another that governed appellate review of federal sentencing.[13] As a consequence of the Court's supposition about what the enacting Congress would have intended about a problem that it neither anticipated nor addressed when legislating, the Federal Sentencing Guidelines are now nonbinding even in the thousands upon thousands of cases in which binding Guidelines pose no constitutional problem.

If severability doctrine is threatening because of the significant discretion it vests in courts to invalidate a statute beyond just those uses and provisions that are unconstitutional, why do we have it? Mostly, it seems, because of the belief that we cannot do without it. Seen or unseen, severability doctrine is omnipresent in judicial review as currently understood. According to existing case law and scholarship, every holding of partial unconstitutionality that does not lead to total invalidation necessarily rests on severability, implicitly if not explicitly. Indeed, according to Michael Dorf, "no workable system of judicial review could function without a large role for severability."[14]

---

[12] *See, e.g.*, Pollock v. Farmers' Loan & Trust Co., 158 U.S. 601, 635–37 (1895) (invalidating as inseparable all income tax provisions of federal tax statute after determining that provisions taxing income derived from real and personal property were unconstitutional).

[13] *Booker*, 543 U.S. at 259 ("[W]e must sever and excise two specific statutory provisions: the provision that requires sentencing courts to impose a sentence within the applicable Guidelines range . . . and the provision that sets forth standards of review on appeal . . . .") (citation omitted).

[14] Dorf, *supra* note 8, at 370.

Operating on the assumption that severability is inevitable in one form or another, scholars have leveled many criticisms at its doctrinal particulars and offered proposals for reform. But nothing has changed. Modern severability doctrine exists in much the same form today as when it emerged in the mid-nineteenth century.

This Article changes the terms of the problem by demonstrating that the counterfactual speculation required by modern severability doctrine is *not* necessary to a workable system of judicial review. While there must be a means of specifying the scope of invalidity flowing from a holding of unconstitutionality, modern severability doctrine is not the only such means.

Although the history of severability has already been written,[15] there is a notable gap in doctrinal history: There exists no sustained scholarly treatment of judicial responses to partially unconstitutional statutes prior to the emergence of legislative intent–based severability analysis in the 1850s. This Article begins to fill that gap by recovering the operative approach to partial unconstitutionality in the first several decades of judicial review following *Marbury*. The evidence that emerges disproves the prevailing notion that judicial review without modern severability doctrine is unworkable. That widely accepted notion rests on an anachronistic attribution of modern severability doctrine to Chief Justice Marshall, his immediate successors on the federal bench, and their state-court contemporaries. Whatever else may be said about the first several decades of post-*Marbury* judicial review in the United States, it was workable. And it worked without modern hypothetical-intent-based severability doctrine.

Judicial review can work that way again today. Most fundamentally, it is necessary to change the foundational metaphor that structures thinking about judicial review—a change from excision to displacement. The familiar excision-based approach to judicial review implies that a court has the power to eliminate unconstitutional provisions by a process of subtraction. In contrast, under a displacement-based approach, a court does not excise anything from a statute but instead determines the extent to which superior law displaces inferior law in resolving the particular case before it.

---

[15] *See* Mark L. Movsesian, *Severability in Statutes and Contracts*, 30 GA. L. REV. 41, 66–73 (1995) (discussing improper importation of contract law concepts into statutory severability analysis); John Copeland Nagle, *Severability*, 72 N.C. L. REV. 203, 210–25 (1993) (tracing development of Supreme Court jurisprudence on severability doctrine); Michael D. Shumsky, *Severability, Inseverability, and the Rule of Law*, 41 HARV. J. ON LEGIS. 227, 240–43 (2004) (describing Supreme Court's reassertion of severability doctrine in historical context of New Deal); Robert L. Stern, *Separability and Separability Clauses in the Supreme Court*, 51 HARV. L. REV. 76, 79–82 (1937) (surveying history and evolution of severability doctrine).

After changing the foundational metaphor, a question remains about the scope of displacement. First principles dictate that superior law can displace no *less* of the inferior law than is repugnant to it, but what about *more*? The proposed approach instructs courts not to infer invalidity beyond unconstitutionality. The goal is for the enforceable law resulting from an exercise of judicial review to be traceable to a combination of constitutional command and legislative provision, rather than judicial supposition about what the legislature would have wanted.

For too long, judicial rewriting through severability doctrine has been accepted as inevitable. This perceived inevitability has kept the debate centered on how courts should rewrite statutes, bypassing entirely the antecedent question of whether they should do so.

This Article proposes a restoration of review without rewriting. It proceeds in three parts. Part I explicates existing doctrine and its flaws. Part II explains how courts in the United States dealt with partially unconstitutional statutes before the emergence of legislative intent–based severability doctrine in the 1850s. Part III expounds a reconstructed version of the earlier approach to partial unconstitutionality that can be implemented in constitutional adjudication in present times.

I
EXCISION

This Part lays the groundwork for the historical analysis in Part II and the doctrinal reconstruction in Part III. This Part first describes the function of severability doctrine in existing law and its fit in the prevailing excision-based approach to judicial review. It then explores the problems with severability doctrine and the approach to judicial review in which it is embedded.

A. The Nature of Severability Doctrine

Severability doctrine governs the decision whether to sever a partially unconstitutional statute. It is one tool for implementing the Supreme Court's general approach to constitutional remedies: to "try to limit the solution to the problem."[16] When the constitutional problem does not inhere in all of a statute's applications or provisions, a reviewing court should, if otherwise appropriate, "enjoin only the unconstitutional applications of a statute while leaving other applications in force, or . . . sever its problematic portions while leaving the

[16] Ayotte v. Planned Parenthood of N. New England, 546 U.S. 320, 328 (2006).

remainder intact."[17] These forms of less-than-total invalidation are typically (if awkwardly) referred to as "application severability" and "text severability" respectively.[18]

Severability doctrine sets legislative intent as a constraint on when courts can "limit the solution to the problem."[19] If the legislature would not have enacted the unconstitutional law without its unconstitutional provisions or applications, the doctrine calls for courts to invalidate the legislation in its entirety.[20] The phrasing of the doctrinal test and the inherently counterfactual nature of the problem reveal that the required inquiry into legislative intent is unlike interpretive inquiries that aim to uncover what the legislature actually provided for in its legislation. Severability doctrine asks what the legislature would have done, not what the legislature actually did.[21]

A legislature may try to control the severability determination by expressing its intent in a severability or inseverability clause.[22] Although one might expect that such a clause unambiguously conveying legislative intent would be dispositive, severability doctrine provides that such clauses are neither necessary nor sufficient to control the judicial determination of severability.[23] Under federal law, "the inclusion of [a severability] clause creates a presumption that Congress did not intend the validity of the statute in question to

---

[17] *Id.* (citation omitted).

[18] *See, e.g.*, Gillian E. Metzger, *Facial Challenges and Federalism*, 105 COLUM. L. REV. 873, 885 & n.52 (2005) (distinguishing "application severability" from "text severability" and identifying sources that discuss them); Adrian Vermuele, *Saving Constructions*, 85 GEO. L.J. 1945, 1950 n.26 (1997) ("[S]everability problems arise not only with respect to different sections, clauses or provisions of a statute, but also with respect to applications of a particular statutory provision when some (but not all) of those applications are unconstitutional.").

[19] *Ayotte*, 546 U.S. at 328.

[20] *See, e.g.*, United States v. Booker, 543 U.S. 220, 249 (2005) (opting not to engraft constitutional requirements onto existing sentencing scheme on ground that Congress would have preferred total invalidation).

[21] *Id.* at 257–58.

[22] *See, e.g.*, 2 U.S.C. § 454 (2006) ("If any provision of this Act, or the application thereof to any person or circumstance, is held invalid, the validity of the remainder of the Act and the application of such provision to other persons and circumstances shall not be affected thereby.").

[23] *See, e.g.*, Dorchy v. Kansas, 264 U.S. 286, 290 (1924) (describing severability clause as "[an] aid in determining [legislative] intent . . . not an inexorable command"); Shumsky, *supra* note 15, at 230 ("Despite the unambiguous command of severability and inseverability clauses, the Court has repeatedly held that they create only a rebuttable presumption that guides—but does not control—a reviewing court's severability determination."). For a general discussion of inseverability clauses, see Israel E. Friedman, Note, *Inseverability Clauses in Statutes*, 64 U. CHI. L. REV. 903, 907–09 (1997).

depend on the validity of the constitutionally offensive provision," but it does not dictate the outcome.[24]

Because severability depends on legislative intent, "questions about the meaning and thus the separability of state statutes are primarily questions of state law," while "[t]he separability of a federal statute is . . . a purely federal issue."[25] Because the precise nature of the severability analysis to be performed depends on the type of law being analyzed (i.e., federal or state law), doctrinal reform of severability doctrine cannot be performed in one fell swoop—the federal system and each state system must be addressed individually. Yet it is still reasonable to generalize about severability doctrine using federal law and to propose a solution for implementation in the federal and state systems alike because state doctrinal systems closely approximate federal doctrine.[26]

B. The Excision-Based Approach to Judicial Review

Severability doctrine is currently nestled within an excision-based approach to judicial review. Within this approach, severability is the exclusive doctrinal tool for dealing with the problem of partial unconstitutionality. But while partial unconstitutionality is ubiquitous, explicit discussions of severability are not. Courts and scholars explain the absence of *explicit* severability reasoning from cases involving partial unconstitutionality by attributing *implicit* severability analysis to those decisions.[27] These attributions of implicit severability rest on an excision-based understanding of judicial review.

According to this understanding, a court facing a partially unconstitutional statute proceeds in two steps. First, the court must determine whether it is possible to "sever" the unconstitutional parts from the constitutional parts. Second, the court must "excise" any unconstitutional parts that are severable, so that the constitutional remainder

---

[24] Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 686 (1987).

[25] RICHARD H. FALLON, JR., DANIEL J. MELTZER & DAVID L. SHAPIRO, HART & WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 182–83 (5th ed. 2003).

[26] *See, e.g.*, Michael C. Dorf, *Facial Challenges to State and Federal Statutes*, 46 STAN. L. REV. 235, 285 (1994) (noting that federal and state severability doctrines are "remarkably uniform" in practice).

[27] *See* United States v. Booker, 543 U.S. 220, 320–32 (2005) (Thomas, J., dissenting) (describing cases in which Supreme Court "disposes of as-applied challenges to a statute by simply invalidating particular applications of the statute, without saying anything at all about severability"); United States v. Nat'l Treasury Employees Union, 513 U.S. 454, 488 (1995) (O'Connor, J., concurring in part and dissenting in part) (describing precedents that held certain statutory applications unconstitutional as "having involved implied severability"); United States v. Grace, 461 U.S. 171, 184 (1983) (Marshall, J., concurring in part and dissenting in part) (describing judgment of partial unconstitutionality as "excising" unconstitutional applications from challenged statute).

can be enforced going forward.[28] If severance and excision are appropriate, the result is conceptualized as a new statute that consists of the old statute "minus" its unconstitutional parts.[29]

Scholars use this theory of excision-based severability doctrine to explain not only every instance of partial invalidation, but also every instance of as-applied adjudication, whether or not it leads to invalidation. A paradigm case is *Yazoo & Mississippi Railroad v. Jackson Vinegar Co.*,[30] in which the Supreme Court rejected a challenge to a Mississippi statute requiring prompt settlement of certain low-dollar-value claims for lost or damaged goods.[31] The Court responded to the railroad's constitutional challenge that the statute improperly required settlement of even false or frivolous claims by declining to adjudicate the constitutionality of the statute itself or even as applied to other cases.[32] The Court instead considered the statute only as applied to the particular claim at issue in the case at hand: The railroad counsel's concession that the underlying claim for a damaged shipment of vinegar bottles was neither false nor frivolous disposed of the case.[33] Said the Court: "It suffices . . . to hold that, as applied to cases like the present, the statute is valid."[34]

Richard Fallon argues that the Court's refusal in *Yazoo* to consider the railroad's "facial" challenge is based on "an implicit assumption that statutory rules are reducible to what might be characterized as statutory 'sub-rules.'"[35] That is, the rule "settle all claims" can be divided into the subrules (i) "settle all valid and non-exorbitant claims," and (ii) "settle all frivolous and excessive claims."[36] With

---

[28] *See, e.g.*, *Booker*, 543 U.S. at 258 ("We now turn to the question of *which* portions of the sentencing statute we must sever and excise as inconsistent with the Court's constitutional requirement.").

[29] *See* Dorf, *supra* note 26, at 238 ("[A] statute that has unconstitutional applications cannot be constitutionally applied to anyone, even to those whose conduct is not constitutionally privileged, unless the court can sever the unconstitutional applications of the statute from the constitutionally permitted ones."); Metzger, *supra* note 18, at 887–88 ("If unconstitutional applications are not severed, the statute cannot be applied to any litigant, even one making no claim of constitutional protection for her conduct. On the other hand, if unconstitutional applications of a statute can be severed, refusing to apply the statute to conduct that is not constitutionally protected becomes unjustified.").

[30] 226 U.S. 217 (1912).

[31] *Id.* at 218–20.

[32] *Id.* at 219–20.

[33] *Id.* at 219.

[34] *Id.* at 219–20.

[35] FALLON ET AL., *supra* note 25, at 182. According to Fallon, the question in a case like *Yazoo* is "whether a statutory provision that does not on its face reflect divisible linguistic units—such as a requirement that railroads must settle 'all claims'—can nonetheless be severed into valid and invalid elements." *Id.*

[36] *Id.*

judicial review understood to be focused on the constitutional validity of subrules, severability doctrine enables as-applied adjudication by providing the mechanism for subrule (i) to survive even if subrule (ii) is invalid.[37]

The reason to attribute implicit severance to as-applied adjudication is to reconcile the practice "with what Henry Monaghan has termed the 'valid rule requirement'—the notion that everyone has a personal constitutional right not to be subjected to governmental sanctions except pursuant to a constitutionally valid rule of law."[38] Michael Dorf explains the logic of this reconciliation, attributing to the *Yazoo* Court the implicit reasoning that "*if* the statute has unconstitutional applications, they are severable from the constitutional applications."[39]

"Excision" is, of course, only a metaphor. There is no judicial Exacto knife that courts use to excise words from the statute books. When a court holds part of a statute unconstitutional, it issues a judgment saying so (and, in some cases, an injunction against its future enforcement). By virtue of precedent and preclusion, this judgment and the reasoning in support of it prevent the unconstitutional part of the statute from having legal effect going forward.[40] Nothing about the actual text of the statute changes as a direct consequence of judicial action.

When a judgment and its supporting reasoning prevent the unconstitutional parts of a statute from having legal effect in all circumstances, it is as if the unconstitutional parts no longer existed, as if the court had excised the unconstitutional parts. But, again, this is just a manner of speaking. Because the court does not do anything to the statute itself, the statute persists as law that can spring back into effect if the constitutional barrier to its enforcement is somehow removed. If constitutional doctrine changes, then no further legislative action is

---

[37] *See id.* ("If the statute is viewed as comprising a number of sub-rules, it becomes comprehensible that sub-rule (i) could survive even if sub-rule (ii) were constitutionally invalid and had to be severed.").

[38] Fallon, *supra* note 6, at 1331 (internal citation omitted).

[39] Dorf, *supra* note 26, at 249. This attribution is necessary to reconcile the Court's conclusion with the possibility that the statute had unconstitutional applications because, "[i]f the statute requiring railroads to settle all claims promptly is an unconstitutional exercise of the state's authority, then, one would suppose, no railroad should be judged by it, even if the railroad's conduct merits no constitutional protection." *Id.* at 243.

[40] *See* Fallon, *supra* note 6, at 1339 ("A court has no power to remove a law from the statute books. When a court rules that a statute is invalid—whether as applied, in part, or on its face—the legal force of its decision resides in doctrines of claim and issue preclusion and of precedent."); David L. Shapiro, *State Courts and Federal Declaratory Judgments*, 74 NW. U. L. REV. 759, 767 (1979) ("No matter what language is used in a judicial opinion, a federal court cannot repeal a duly enacted statute of any legislative authority.").

necessary for "excised" portions of the statute to regain their legal effect.[41]

Although excision is only a metaphor, it is a powerful one that underwrites the perceived need for severability doctrine. Michael Dorf attempts to establish the central role of severability doctrine to constitutional adjudication by demonstrating the unworkability of a general rule of nonseverability.[42] If courts were to apply a rule of nonseverability, Dorf asks, "[W]ould the invalidation of some snippet of text imply the invalidation of everything else that was enacted as part of the same omnibus bill as that snippet? Would it imply the invalidation of the provision in which the snippet was embedded at the time of the challenge?"[43] Indeed, Dorf presses, "[w]hy just the provision rather than the Code section, the Code title, or the entire U.S. Code itself?"[44] Dorf concludes that "[b]ecause the more extreme of these options are plainly implausible, courts never face a choice of *whether* to sever invalid provisions or applications from valid ones, but instead must always decide *how much* to sever."[45] The argument, in essence, is that severability is inevitable because nonseverability is unworkable.

In sum, severability doctrine is viewed as the exclusive tool for dealing with partial unconstitutionality under the prevailing excision-based understanding of judicial review. According to this understanding, "a statute that has unconstitutional applications cannot be constitutionally applied to anyone, even to those whose conduct is not constitutionally privileged, unless the court can sever the unconstitutional applications of the statute from the constitutionally permitted ones."[46] Moreover, the argument goes, every occasion of invalidation results in judicial lawmaking through subtraction, and it has been this way since *Marbury*.[47] In this excision-based view, severance is salvific: It saves the constitutional remainder of an unconstitutional law by

---

41 It may, however, be necessary for a party to prior litigation holding the statute unconstitutional to seek modification of an injunction pursuant to Rule 60(b). FED. R. CIV. P. 60(b)(5); *see also Agostini v. Felton*, 521 U.S. 203, 239–40 (1997) (authorizing relief from permanent injunction limiting local educational officials' use of federal funds based on significant changes in Establishment Clause doctrine).

42 Dorf, *supra* note 8, at 370.

43 *Id.*

44 *Id.*

45 *Id.*; *see also id.* ("[Without severability,] any judicial decision finding any law unconstitutional, on its face or as applied, would call into question the entire legal code.").

46 Dorf, *supra* note 26, at 238.

47 *See id.* at 250 (describing Supreme Court's opinion in *Marbury* as applying "the same presumption of severability" as in *Yazoo*); *cf.* Fallon, *supra* note 6, at 1337 n.91 ("*Marbury* provides an especially nice illustration of the extent to which constitutional practice presupposes that statutory rules—understood as linguistic units—can and should be treated as comprising severable subrules.").

treating the unconstitutional portion "as if it were a distinct unconstitutional provision."[48]

The excision-based understanding of judicial review provides a powerful, internally coherent account of constitutional adjudication in the United States. This understanding does have one descriptive weakness: It requires attribution of implicit severance on a massive scale. But this weakness is minor—such attribution does not contradict anything the Court has done and is consistent with the understanding of at least some Justices.[49]

The principal problem with excision is not descriptive but prescriptive: Excision requires deployment of a destructive doctrine that is subject to manipulation because of the counterfactual speculation that it requires.

C. Problems with Severability and Excision

Despite its potency, severability doctrine often evades scrutiny by remaining in the shadows of the substantive constitutional rulings that occasion its application. Those who have trained their lights on the doctrine, however, have been almost uniformly critical of what they have seen.

Robert Stern wrote the seminal article on severability in 1937.[50] Frustrated by the Supreme Court's complete invalidation of partially unconstitutional New Deal economic legislation, Stern sought to bring order to a doctrine that he diagnosed as riddled with contradiction.[51] Yet Stern could find no organizing principles behind the doctrine, and he concluded that the Court's severability reasoning simply rationalized results reached on other grounds (such as preferred outcomes).[52] This criticism has never satisfactorily been answered. But as the Court showed greater receptivity to New Deal legislation as a matter of substantive constitutional law, the pressure on severability doctrine abated.[53]

---

[48] Dorf, *supra* note 26, at 250.

[49] *See supra* note 27.

[50] Stern, *supra* note 15.

[51] *Id.* at 78 ("Only if the apparent inconsistencies in Supreme Court decisions [dealing with the severability of statutes] are exposed and explained, as they have not been by the Court itself, can a conscious effort be made to formulate understandable and sensible principles for the future.").

[52] *Id.* at 101–02 ("[T]he Court avails itself of one [severability] formula or another in order to justify results which seem to it to be desirable for other reasons.").

[53] *See* Shumsky, *supra* note 15, at 240 ("Since 1936, . . . severability questions seem to have mostly faded into the background—perhaps as a consequence of the Court's willingness to accept the constitutionality of most New Deal-style regulatory legislation.").

While the shift in substantive law diverted scholarly attention to other matters, the doctrine stayed the same for the next several decades, warts and all. Evaluating severability doctrine in the mid-1990s, Mark Movsesian observed:

> [Severability doctrine] has drawn criticism on almost every conceivable basis. Commentators have condemned [it] as too malleable and as too rigid; as encouraging judicial overreaching and as encouraging judicial abdication. They have criticized the doctrine's reliance on legislative intent and its disregard of legislative intent; its excessive attention to political concerns and its inattention to political concerns . . . .[54]

Though these criticisms obviously reflect differing evaluations, they share the common characteristic of dissatisfaction with existing severability doctrine. Recent scholarship has continued in this same critical vein.[55]

The most recent spur to critical evaluation of severability doctrine is the Court's audacious use of that doctrine in *United States v. Booker*.[56] Justice Breyer's remedial opinion for the Court used severability doctrine to render the Federal Sentencing Guidelines nonmandatory by "severing" and "excising" two statutory provisions from the Sentencing Act.[57] The effect was to make the Guidelines advisory in *all* federal sentencing even though mandatory Guidelines would be

---

[54] Movsesian, *supra* note 15, at 41–42 (internal citations omitted).

[55] *See, e.g.*, David H. Gans, *Severability as Judicial Lawmaking*, 76 GEO. WASH. L. REV. 639, 663 (2008) ("[Severability doctrine] gives courts extensive power to rewrite statutes . . . in a way that makes after-the-fact legislative correction unlikely[,] . . . warps legislatures' incentives to obey constitutional norms ex ante[,] [and] allows courts to make vague law without thinking about it."); Nagle, *supra* note 15, at 225 ("The confusion surrounding presumptions and the absence of a consistent effort to explain how severability fits within broader theories of judicial review and statutory construction has left all of the various tests used over the years unanchored by a principled approach."); Emily L. Sherwin, *Rules and Judicial Review*, *in* 6 LEGAL THEORY 299, 308 (2000) (arguing that existing doctrine inadequately accounts for "the effect of severance and related practices in establishing new rules, and the relative competence of courts and legislatures to perform the rulemaking function").

[56] 543 U.S. 220, 259 (2005).

[57] *Id.* Justice Breyer first inquired whether Congress would have preferred federal sentencing with the Sixth Amendment jury requirement engrafted onto the current system, or instead would have preferred no determinate sentencing scheme at all. *Id.* at 249. He concluded that "Congress would likely have preferred the total invalidation of the Act to an Act with the Court's Sixth Amendment requirement engrafted onto it." *Id.* Next, he asked whether some remedy short of complete invalidation would both solve the constitutional problem and also be consistent with legislative intent, concluding that "Congress would likely have preferred the excision of some of the Act, namely the Act's mandatory language, to the invalidation of the entire Act." *Id.* The Court accordingly "sever[ed] and excise[d]" the provision that made the Guidelines mandatory, along with intertwined provisions relating to appeal. *Id.* at 259.

unconstitutional in only *some* federal sentencing.[58] The *Booker* remedy is a form of statutory *cy pres*—the judicial crafting of a new sentencing scheme that conforms as far as possible to hypothetical legislative intent in light of the constitutional barrier to the implementation of the sentencing scheme that Congress actually enacted. As the Court must have recognized, and as the cases following *Booker* have borne out, the *Booker* remedy committed the federal judiciary to working out in a common law–like fashion the operational details of the new advisory Guidelines scheme wrought by the Court's severance and excision.

Justice Breyer's opinion for the Court denied the dissenters' charges that it had created a new kind of severance analysis.[59] In a thorough analysis of the *Booker* remedy, Gillian Metzger similarly explains that, although "reasonable minds can[ ] and do[ ] differ" about whether "Congress would have preferred invalidation of [only] the Act's mandatory provisions to its total invalidation," these differences speak only to this specific application of severability doctrine, not to the legitimacy of using severability doctrine in general.[60] Metzger concludes that, "accepting arguendo the majority's reading of the Sentencing Act, its remedial approach of facial invalidation was perfectly legitimate."[61]

Evaluated through this lens, the *Booker* remedy reveals how modern severability doctrine provides for a form of "backdoor" facial invalidation that depends not on constitutional command but on judicially perceived counterfactual legislative intent. When a court uses application inseverability to invalidate all applications of a provision that is only unconstitutional in some applications, the effect is equivalent to a blanket "excision" of text that is only sometimes unconstitutional. And barring either a change in the underlying constitutional rule or new legislation, the result is that the *sometimes* unconstitutional text may *never* be judicially enforced.

---

[58] *Id.* at 266–67. The remedy of making the Guidelines advisory in all federal sentencing was not argued for by any party or amicus curiae.

[59] *Compare id.* at 247 (majority opinion), *with id.* at 283 (Stevens, J., dissenting) ("The precedent on which the Court relies is scant indeed."), *and id.* at 325 (Thomas, J., dissenting) (arguing that majority's remedial application of severability "constitute[s] legislation beyond [the Court's] judicial power").

[60] Metzger, *supra* note 18, at 892 (quoting *Booker*, 543 U.S. at 248). Metzger grounds her defense of the *Booker* remedy in the principle of "application severability," arguing that "if a provision's unconstitutional applications cannot be severed or construed away, then . . . [the] provision cannot be constitutionally applied at all . . . ." *Id.* at 892. According to Metzger, "application inseverability" supplies the source of judicial authority to invalidate a statutory provision that is not facially unconstitutional. *Id.*

[61] *Id.*

This analysis of the *Booker* remedy also illustrates just how much lawmaking authority severability doctrine grants to a court. By reaching determinations about counterfactual legislative intent regarding a combination of applications and provisions, a court reviewing a partially unconstitutional statute can expand the scope of its invalidation as widely or as narrowly as it discerns to be consistent with hypothesized legislative intent. This authority to "excise" and "rewrite" is effectively discretionary because the legislative intent test is almost always indeterminate. The Court claims to discern what Congress *would have intended* had Congress known that its actual legislation was unconstitutional as written. In *Booker*, the Court's conclusion regarding "what Congress would have intended in light of the Court's constitutional holding"[62] looked very much like what the Court would have legislated had it possessed the authority to make sentencing policy as a legislator.

It should come as no surprise that the Court's conclusion in *Booker* about what Congress *would* have wanted lines up closely with what five Justices think Congress *should* have wanted. And it is impossible to assail the outcome with evidence of what the enacting Congress *did* want because that Congress expressed no actual legislative intent regarding mandatory sentencing under the *Apprendi*/*Blakely* line of cases.[63]

This is not an isolated problem. Severability doctrine requires focus on hypothetical intent because there often is no actual,

---

[62] *Booker*, 543 U.S. at 246 (internal quotations omitted).

[63] Determinate federal sentencing based on binding guidelines promulgated by the United States Sentencing Commission resulted from the Sentencing Reform Act of 1984. *See* Mistretta v. United States, 488 U.S. 361, 367–68 (1989) (describing revisions to federal sentencing in 1984 Act). The Fifth and Sixth Amendment limitations on judicial factfinding that doomed the Guidelines in *Booker* began to emerge in the mid-to-late 1990s. *See* Jones v. United States, 526 U.S. 227, 243 n.6 (1999) (arguing that under Fifth and Sixth Amendment limitations, "any fact (other than prior conviction) that increases the maximum penalty for a crime must be charged in an indictment, submitted to a jury, and proven beyond a reasonable doubt"); Douglas A. Berman, *Examining the* Blakely *Earthquake and Its Aftershocks*, 16 FED. SENT'G REP. 307, 307 (2004) ("[T]he constitutional ground under sentencing reform had been rumbling for some time before the *Blakely* earthquake hit."). These limitations were given serious bite in Apprendi v. New Jersey, 530 U.S. 466 (2000), in which the Court held unconstitutional a New Jersey hate-crime statute that provided for a sentencing enhancement based on judicial factfinding. But it was not until Blakely v. Washington, 542 U.S. 296 (2004), that the principle underlying *Apprendi* was understood as a direct threat to the federal guidelines system. *See* Berman, *supra*, at 308 (explaining that Court's holding in *Blakely* came as surprise because "most observers believed [*Blakely*] was to serve as final confirmation that the *Apprendi* decision would not radically transform modern sentencing practices"). There is no way that the Congress that created the Guidelines system two decades earlier could have formed an intent about how the system should operate in light of the constitutional requirements identified in *Blakely*.

expressed legislative intent to be found.[64] The hypothetical legislative intent test gets around this absence of any actual legislative intent to discern but does so by posing a question whose answer often calls for rank speculation.[65] The inquiry "can perhaps be given some content by reference to legislative purpose, or to the 'enterprise' to which the statute belongs. But it can easily deteriorate into a question of what ought to happen, in which case 'legislative intent' adds nothing."[66]

There is another problem. Suppose the hypothetical legislative intent approach yields an as-close-as-you-can-get-to-determinative conclusion that the legislature would not have enacted the legislation without its unconstitutional parts. Severability doctrine would then require total invalidation, even if the reason the legislature would not have enacted the legislation has nothing to do with either the constitutional values or legitimate legislative objectives at issue. Suppose, for example, that swing votes for an important piece of reform legislation are bought through appropriations to a few legislators' states. Suppose, further, that the appropriations are enacted as part of the same omnibus bill but are otherwise unrelated to the subject matter of the reform. If those appropriations are later held unconstitutional, and the legislature would not have enacted the statute without those vote-buying payoffs, severability doctrine would require invalidation of the reform legislation.

A concrete illustration of this potential problem can be found in the path through the federal courts followed by the Bipartisan Campaign Finance Reform Act ("BCRA"). In 2003, the BCRA survived a facial challenge in *McConnell v. FEC*.[67] Five years later, the Court held that a BCRA provision whose constitutionality the Court did not adjudicate in *McConnell*—the so-called "Millionaires'

---

[64] As Emily Sherwin has observed, "severability questions are triggered by unplanned statutory failures." Sherwin, *supra* note 55, at 304; *see also* Stern, *supra* note 15, at 98 ("Difficult problems of statutory construction generally arise because the legislature has not thought of the particular situation which has come before the Court, and accordingly had no real intention as to how the law should be construed with respect to it."). The generalization about unplanned statutory failures does not always hold true, particularly with respect to legislation in fertile areas of constitutional litigation. But it holds often enough to assign the absence of forethought as one reason for the absence of actual intent.

[65] *See, e.g.*, Sherwin, *supra* note 55, at 305 ("The difficulty here is that hypothetical intent is quite speculative.").

[66] *Id.* These criticisms are aimed not at the consideration of legislative intent in general, but at the counterfactual speculation required by modern severability doctrine in particular.

[67] 540 U.S. 93 (2003) (upholding against facial First Amendment challenge various provisions of Bipartisan Campaign Finance Reform Act of 2002, Pub. L. No. 107-155, 116 Stat. 81 (codified in scattered sections of 2 U.S.C.) (regulating soft-money contributions and "electioneering communications")).

Amendment"—violated the First Amendment.[68] Suppose it can be shown that Congress would not have enacted the BCRA absent the incumbency protection provided by the Millionaires' Amendment.[69] If that supposition is correct, then the BCRA must fall in its entirety. While totally invalidating the BCRA based on severability doctrine after upholding it against a facial challenge might seem absurd, that is exactly what adherence to the hypothetical legislative intent test would require.[70] The alternative is to fight the hypothetical and dispute whether Congress would have enacted the BCRA without the Millionaires' Amendment. But that move only returns focus to the indeterminacy of the inquiry: Can anyone really know whether the BCRA would have passed without the Millionaires' Amendment? Should the continued existence of the entire legislative scheme for federal campaign financing turn on the answer to such a question? Would anyone be surprised if this question were asked of the Supreme Court, and the answers tracked the Justices' personal views about the constitutionality of the BCRA more generally?[71]

Although *Booker* and the BCRA provide particularly vivid illustrations of the sweeping discretionary authority supplied by modern severance doctrine, the problems discussed above are not new, and severability doctrine has remained fundamentally unchanged since it was introduced over a century and a half ago. There is nothing much

---

[68] Davis v. FEC, 128 S. Ct. 2759 (2008). The Millionaires' Amendment relaxed contribution limits for candidates running against opponents who spent significant sums of their own money in support of their candidacy. *See id.* at 2765–67 (describing how "Millionaires' Amendment" provision operated). The Court held that "imposing different contribution and coordinated party expenditure limits on candidates vying for the same seat is antithetical to the First Amendment." *Id.* at 2774.

[69] *But see* Transcript of Oral Argument at 47–48, *Davis*, 554 U.S. ___ (No. 07-340) (Solicitor Gen. Paul Clement) (arguing that Millionaires' Amendment is severable, relying in part on presence of severability clause); 34 Op. FEC 9 (2008) (concluding that personal loan provision in BCRA was severable from unconstitutional Millionaires' Amendment).

[70] A wrinkle in this argument is that some legislators who voted for the Millionaires' Amendment may have recognized the possibility that it was unconstitutional but voted to include it anyway in the belief that the statute's severability clause would prevent its unconstitutionality from bringing down the rest of the statute. Of course, this additional wrinkle just further exposes how difficult the project of discerning hypothetical legislative intent can be.

[71] *Cf.* Stern, *supra* note 15, at 113 ("[T]he judges who disapproved of the law thought it was inseparable, while those who regarded it as constitutional also took a contrary position on the question of separability.").

The Court's recent decision in *Citizens United v. FEC*, 130 S. Ct. 876 (2010), presents another opportunity for speculation about the BCRA's severability. Would Congress have enacted the legislation if it knew that it could not ban independent corporate spending on electioneering communications? What about the ban on electioneering communications as applied to others? Given the stakes of the answers to these questions and the indeterminacy of the inquiry, perhaps the answers will depend on who one asks.

to do about this state of affairs if modern severability doctrine is necessary for "[any] workable system of judicial review."[72] Must we just learn to live with it and tinker at the edges of the doctrine? If so, then observing the doctrine's unyielding endurance to sustained criticism would be an occasion for lamentation, but nothing more.

The truth is that modern severability doctrine is not inevitable. We can have a functional system of judicial review without it. Indeed, the system would function better without it. The next Part aims to puncture the inevitability of modern severability doctrine through historical analysis. Part III then fills the conceptual space opened up by the elimination of modern severability doctrine.

II
Doctrinal History

This Part presents the results of research into the original approach to partial unconstitutionality. It first identifies the conceptual foundations of the approach and then describes how the background principles for dealing with partial unconstitutionality operated across a range of areas of substantive constitutional law. This Part closes with a discussion of why the original approach to partial unconstitutionality has not been previously noticed.

A. Foundations of Repugnancy Displacement

The recovery project begins with a fresh reading of Alexander Hamilton's exposition in Federalist No. 78, which Philip Hamburger describes as "an unusually elegant illustration of how lawyers tended to perceive the authority of the judges to expound the law, including the U.S. Constitution."[73] According to Hamilton, "The judiciary . . . has no influence over either the sword or the purse; no direction either of the strength or of the wealth of the society, and can take no active resolution whatever."[74] The power of the judiciary is to enter judgments in cases.[75] In reaching judgments, courts possess the authority "to pronounce legislative acts void, because contrary to the constitution . . . ."[76] This authority derives from the Constitution's status as fundamental law and from judges' obligation "to regulate

---

[72] Dorf, *supra* note 8, at 370.

[73] Philip Hamburger, Law and Judicial Duty 552 (2008).

[74] The Federalist No. 78, at 378 (Alexander Hamilton) (Terence Ball ed., 2003).

[75] *See id.* ("It may truly be said to have neither Force nor Will, but merely *judgment*; and must ultimately depend upon the aid of the executive arm even for the efficacy of its judgments.") (emphasis added).

[76] *Id.* at 379.

their decisions by the fundamental laws, rather than by those which are not fundamental."[77]

The practice that Hamilton defended in Federalist No. 78 is not the same as the modern conception of "judicial review," in which courts "strike down" legislation. Rather, what we now call judicial review consisted of a refusal to give a statute effect as operative law in resolving a case. Hamilton described the practice of refusing to enforce an unconstitutional statute as an "exercise of judicial discretion in determining between two contradictory laws . . . ."[78] Because both the Constitution and statutes are law, it belongs to the courts to ascertain the meaning of each, and "[i]f there should happen to be an irreconcileable variance between the two, that which has the superior obligation and validity ought of course to be preferred . . . ."[79] Hamilton recognized that two laws could be totally or only partially contradictory,[80] but he thought the same principle governed both cases: Courts should reconcile the laws to the extent possible and, where this cannot be done, "give effect to one, in exclusion of the other."[81] The operative principle for Hamilton was not excision but displacement.

According to Mary Sarah Bilder, "[w]hat contemporary scholars and the public call 'judicial review' originally had no such name. Indeed, until the early twentieth century, commentators identified it descriptively as courts voiding legislation repugnant to a constitution."[82] Similarly, Philip Hamburger explains that what is now called judicial review was an implication of the nature of constitutions as law and the judicial duty "to decide in accord with the law of the land, including the constitution."[83]

Other scholars have described American judicial review in its origins as requiring that Article III courts "*refuse to enforce* legislation that violates the Constitution."[84] Much of the evidence that these scholars rely on is also evidence of a shared presupposition that courts

---

[77] *Id.* at 380.

[78] *Id.*

[79] *Id.* at 379–80.

[80] *See id.* at 380 ("It not uncommonly happens, that there are two statutes existing at one time, *clashing in whole or in part* with each other, and neither of them containing any repealing clause or expression.") (emphasis added).

[81] *Id.*

[82] Mary Sarah Bilder, *Idea or Practice: A Brief Historiography of Judicial Review*, 20 J. POL'Y HIST. 6, 8 (2008).

[83] HAMBURGER, *supra* note 73, at 17.

[84] Saikrishna B. Prakash & John C. Yoo, *The Origins of Judicial Review*, 70 U. CHI. L. REV. 887, 914 (emphasis added); *see also id.* at 916 ("[V]indicating the people's choice of a limited Constitution requires judges to refuse to enforce unconstitutional statutes."); *id.* at 955 ("Throughout the United States, dozens of speakers and writers made clear that

could and would engage in case-by-case displacement of unconstitutional statutes.[85] The displacement takes place by virtue of a rule of priority that requires application of superior law over inferior law to the extent there is a conflict between the two.[86]

Viewed in this light, it is anachronistic to view *Marbury* as a case of implicit severance and excision. It is, instead, an instance of displacement. The statutory grant of authority for the Supreme Court to issue writs of mandamus in the exercise of its original jurisdiction went beyond the Article III grant of original jurisdiction to the Supreme Court. As inferior law, the statutory grant was therefore unconstitutional to the extent that it went beyond Article III, but no further. The Judiciary Act of 1789 was still "law" to the extent that it was not "void." The operative concept of judicial review here was one in which the Court, in disposing of the case before it, was obliged to apply the superior law of the Constitution in preference to conflicting, inferior statutory law: "If then the courts are to regard the constitution; and the constitution is superior to any ordinary act of the legislature; the constitution, and not such ordinary act, *must govern the case to which they both apply*."[87]

B. Applications

The concept of judicial review as refusal to give effect rather than "striking down" may seem weak. But experience from the first several decades of judicial review shows that the practice could be muscular and effective in ensuring the supremacy of the Constitution over conflicting laws. Operating within an intellectual framework in which judicial review consisted of a refusal to give effect to inferior law that

judges, federal and state, could refuse to enforce legislation that transgressed the Constitution.").

[85] *See* 2 THE RECORDS OF THE FEDERAL CONVENTION OF 1787, at 78 (Max Farrand ed., rev. ed. 1937) (statement of Col. Mason) ("[I]f the Judges were joined in this check on the laws, . . . *they could impede in one case only, the operation of laws*. They could declare an unconstitutional law void.") (emphasis added); BRUTUS XII (1788), *reprinted in* THE FEDERALIST WITH LETTERS OF "BRUTUS" 506, 508 (Terence Ball ed.) ("[T]he courts are vested with the supreme and uncontroulable power, to determine, in all cases that come before them, what the constitution means; *they cannot therefore, execute a law, which, in their judgment, opposes the constitution*, unless we can suppose they can make a superior law give way to an inferior.") (emphasis added). Although most of this evidence speaks to only the shared presupposition of case-by-case displacement of unconstitutional statutes, there are some statements that come closer to speaking to the issue of partial unconstitutionality. For example, in congressional debates over the First Judiciary Act, Elbridge Gerry stated that "[t]he Constitution will undoubtedly be [the courts'] first rule; and *so far as* your laws conform to that, they will attend to them, *but no further*." 1 ANNALS OF CONG. 829 (Joseph Gales ed., 1834) (emphases added).

[86] *See supra* notes 78–81 and accompanying text.

[87] Marbury v. Madison, 5 U.S. (1 Cranch) 137, 178 (1803) (emphasis added).

was repugnant to superior law, federal and state courts were able to vindicate the Contracts Clause,[88] the Commerce Clause,[89] and the Import-Export Clause,[90] among other provisions, without massive displacement of partially unconstitutional state laws.

1. "Not the Terms of the Law But the Effect"

Some of the clearest examples of the original approach to partial unconstitutionality appear in cases invalidating state legislation under the Contracts Clause.[91] The doctrinal test applied in these cases often yielded application-specific holdings of unconstitutionality. These holdings operated in combination with the background principle that a statute may be void in part and valid in part to preserve challenged statutes except insofar as their application would be unconstitutional. That is, invalidity did not extend beyond unconstitutionality, and unconstitutionality was confined to a specific application or set of applications. The output of adjudication was not about the constitutional validity of a statute in the abstract, but rather about the statute's constitutionality with respect to a particular application or class of applications.

Golden v. Prince, an 1814 opinion by Justice Bushrod Washington riding circuit, is illustrative.[92] A debtor invoked a Pennsylvania statute as discharging his debts in defense to suit for payment on a bill of exchange. The creditor responded that the statute was doubly unconstitutional, first because the power to make bankruptcy laws was exclusively federal, and second because the state law impaired the obligation of contracts in violation of the Contracts Clause.[93] Justice Washington ruled for the creditor on both grounds and, in so doing, explicitly differentiated the scope of the resulting statutory invalidity for each holding.[94]

Close attention to the reasoning in this opinion provides three important insights into the operative understanding of partial unconstitutionality. First, the function of judicial review was to determine

---

[88] U.S. CONST. art. I, § 10, cl. 1 ("No State shall . . . pass any . . . Law impairing the Obligation of Contracts . . . .").

[89] U.S. CONST. art. I, § 8, cl. 3 ("The Congress shall have Power . . . To regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes . . . .").

[90] U.S. CONST. art. I, § 10, cl. 2 ("No State shall, without the Consent of the Congress, lay any Imposts or Duties on Imports or Exports, except what may be absolutely necessary for executing [its] inspection Laws . . . .").

[91] U.S. CONST. art. I, § 10, cl. 1.

[92] 10 F. Cas. 542 (C.C.D. Pa. 1814) (No. 5509).

[93] *Id.*

[94] *Id.* at 547.

whether to recognize the challenged law as operative law *in the case before the court*. Justice Washington described the constitutional analysis as examining "whether the law relied upon by the defendant, to bar the present action, is repugnant to the constitution of the United States; and, on that account, is not to be regarded by the court, *in this case*[.]"[95]

Second, statutory law could be unconstitutional on some occasions but not others. This feature of Justice Washington's approach can be seen in the premise with which he began his Contracts Clause analysis: "[A] law may be unconstitutional, and of course void, in relation to particular cases; and yet valid to all intents and purposes, in its application to other cases within the scope of its provisions, but varying from the other in particular circumstances."[96] For instance, a debtor discharge statute is unconstitutional as applied to contracts entered into *before* the statute's enactment but not as applied to contracts entered into *after* the statute's enactment.[97]

Third, some constitutional defects can render a statute unconstitutional in all circumstances. For instance, if the power to enact bankruptcy laws is exclusively federal, and the state law is an exercise of that power, then the state law is unconstitutional in all applications.[98]

With these principles framing his analysis, Washington first found the Pennsylvania statute unconstitutional as applied to the discharge of Prince's debt on the bill of exchange entered into prior to the law's passage.[99] Next, Washington found the statute's bankruptcy legislation fully unconstitutional because the power to make bankruptcy laws was exclusively federal.[100] The scope of statutory invalidity flowing from each flaw differed based on the scope of the invalidating principle of constitutional doctrine: "We are, upon the whole, of opinion, that the law under which the certificate is pleaded, in bar of the action, is *altogether unconstitutional*, for the reason last mentioned [i.e., that the

---

[95] *Id.* at 544 (emphasis added).

[96] *Id.*

[97] *Id.* According to Washington, "a law prospective in its operation, under which a contract afterwards made, may be avoided in a way different from that provided by the parties, would be clearly constitutional." *Id.* By contrast, "if the law act retrospectively, as to other contracts, so as to impair their obligation, the law is invalid; or in milder terms, it affords no rule of decision in these latter cases." *Id.*

[98] *Id.* at 545–47.

[99] *Id.* at 544 ("The question then is, whether a law of a state, which declares that a debtor, by delivering up his estate for the benefit of his creditors, shall be for ever discharged from the payment of his debts, due or contracted before the passage of the law . . . can be set up to bar the right of such creditor to recover his debt, either in a federal or state court? . . . [I]t cannot; because the law is . . . one which, in the case supposed, impairs the obligation of a contract.").

[100] *Id.* at 545–47.

power to enact bankruptcy legislation is exclusively federal]; *and is so in reference to this debt*, for the first reason [i.e., because it impaired the obligation of contracts]."[101]

The Supreme Court confirmed Justice Washington's approach to partial unconstitutionality when it held in *Sturges v. Crowninshield* that the federal power to make uniform bankruptcy laws did not prevent states from enacting their own debtor discharge laws in the absence of a conflicting federal bankruptcy statute.[102] Notwithstanding this holding recognizing state power, the Court held the law in *Sturges* unconstitutional *with respect to the particular debt at issue*.[103] The Court could not apply the statute to the debt at issue because that debt had been contracted for *before* the discharge statute was enacted.[104]

The idea that a statute can be unconstitutional with respect to a particular debt or class of debts, but otherwise valid, rested on the application of a doctrinal test that looked to statutory effects rather than statutory language. The opinions of Justices Trimble, Washington, and Thompson in the Supreme Court's fractured 4-3 decision in *Ogden v. Saunders* reveal how this is so.[105] The analysis in Justice Trimble's opinion[106]—his only signed opinion in a constitu-

---

101 *Id.* at 547.

102 17 U.S. 122, 208 (1819).

103 *Id.* ("[The Act], *so far as it attempts to discharge this defendant from the debt in the declaration mentioned*, is contrary to the constitution of the United States, and . . . the plea is no bar to the action.") (emphasis added). The Court's certificate sending the case back down to the lower court also contains language explicitly delineating the scope of the Court's holding:

> CERTIFICATE. . . . [T]his Court is of opinion, that, since the adoption of the constitution of the United States, a State has authority to pass a bankrupt law, provided such law does not impair the obligation of contracts, within the meaning of the constitution, and provided there be no act of Congress in force to establish a uniform system of bankruptcy, conflicting with such law. This Court is farther of opinion, that the act of New-York, which is pleaded in this case, *so far as it attempts to discharge the contract on which this suit was instituted*, is a law impairing the obligation of contracts within the meaning of the constitution of the United States, and that the plea of the defendant is not a good and sufficient bar of the plaintiff's action.

Id. (emphasis added).

104 *Id.* at 197–98.

105 25 U.S. 213 (1827).

106 According to David P. Currie, "*Ogden* was the only constitutional case in thirty-four years in which Marshall signed a dissent, and he took Story and Duvall with him." DAVID P. CURRIE, THE CONSTITUTION IN THE SUPREME COURT: THE FIRST HUNDRED YEARS 1789–1888, at 151 (1985).

The Court addressed two issues in *Ogden*. The first was left open in *Sturges*—whether a state debtor discharge law violated the Contracts Clause as applied to a contract entered into *after* the statute had been enacted. *See Ogden*, 25 U.S. at 254 (stating that "whether the obligation of a contract is impaired by a State bankrupt or insolvent law, which dis-

tional case[107]—is the most explicit:

> It is not the terms of the law, but its effect, that is inhibited by the constitution. A law may be in part constitutional, and in part unconstitutional. It may, when applied to a given case, produce an effect which is prohibited by the constitution; but it may not, when applied to a case differently circumstanced, produce such prohibited effect. Whether the law under consideration, in its effects and operation upon the contract sued on in this case, be a law impairing the obligation of this contract, is the only necessary inquiry.[108]

Justice Thompson's and Justice Washington's opinions also exhibit two important operative principles: (1) the Constitution is concerned with effects and not words, and (2) a statute may be void in part and valid in part. Wrote Justice Thompson: "It was not denied on the argument, and, I presume, cannot be, but that a law may be void in part and good in part;. . . it may be void, so far as it has a retrospective application to past contracts, and valid, as applied prospectively to future contracts."[109] Similarly, Justice Washington commented: "[W]hich ever way we turn, . . . we are met by this overruling and admitted distinction, between those which operate retrospectively, and those which operate prospectively. In all of them, the law is pronounced to be void in the first class of cases, and not so in the second."[110]

---

charges . . . liability under a contract entered into in that State after the passage of the act" is a question that "has never before been distinctly presented to the consideration of this Court, and decided"). On this issue, the Court ruled that prospective application of a debtor discharge law did not impair the obligation of a contract. *Id.* at 368–69 (Johnson, J.). The second issue related to territorial limits on legislative jurisdiction. On this issue, Justice Johnson joined the three dissenters from the Court's disposition of the first issue and held that the New York debtor discharge law could not constitutionally be applied to the detriment of Saunders, a citizen of Kentucky who had "never voluntarily subjected himself to [New York's] laws, otherwise than by the origin of his contract." *Id.* at 358. Justice Johnson summarized the differential effect of these conclusions on the validity of the New York debtor discharge law: "[A]s between citizens of the same State, a discharge of a bankrupt by the laws of that State, is valid as it affects posterior contracts; [and] as against creditors, citizens of other States, it is invalid as to all contracts." *Id.* at 368–69.

107 CURRIE, *supra* note 106, at 194–95 (1985) ("When we attempt to analyze the work of individual Justices, the most striking fact is that most of Marshall's brethren were nearly invisible. . . . Trimble wrote once, for himself alone, to uphold the prospective bankruptcy law.").

108 *Ogden*, 25 U.S. at 316.

109 *Id.* at 295 (Johnson, J.).

110 *Id.* at 262 (Washington, J.). The Court applied the distinction between retrospective and prospective application when assessing Contracts Clause challenges not only to debtor discharge laws, but also to other types of statutes. In *Bronson v. Kinzie*, for example, the Court held invalid an Illinois mortgage moratorium law that limited the remedies available to a lender as applied to mortgages entered into prior to the law's passage. 42 U.S. (1 How.) 311, 320–22 (1843). The Court's opinion explicitly recognized that the law could be applied constitutionally to mortgages entered into after the statute's enactment. *Id.* at 321.

2. *"To the Extent of Such Collision and Repugnancy"*

The Court did not limit the foregoing principles shaping its approach to partial unconstitutionality to the Contracts Clause alone. When the Marshall Court's Commerce Clause and Import-Export Clause doctrines ran up against the exercise of state powers addressing the same subject matter as federal law, the same principles of partial unconstitutionality greased the gears of the federal machine to keep it running. Two pairs of cases—each consisting of a Marshall opinion followed by an opinion by Chief Justice Lemuel Shaw for the Massachusetts Supreme Judicial Court—illustrate this point.

The first pair of cases, *Gibbons v. Ogden*[111] and *Norris v. Boston*,[112] addresses the distribution of power to regulate interstate commerce. *Gibbons* provided an application-specific framework for analyzing displacement of state law by federal law, which the later decision in *Norris* relied upon in elaborating its approach to the problem of partial unconstitutionality. Rather than decide the momentous question that consumed much of the argument in *Gibbons*—whether the states possessed a concurrent power to regulate interstate commerce—Marshall rested his decision on the actual existing conflict between the federal statute under which Gibbons had obtained his license and the grant of exclusivity resulting from New York state statutes.[113] The Court asked "whether the laws of New-York, as expounded by the highest tribunal of that State, have, *in their application to this case, come into collision* with an act of

Relatedly, in *Woodruff v. Trapnall*, the Court partially invalidated legislation that repealed a provision of the charter of the Bank of Arkansas providing that its notes would be accepted in payment of public debts. 51 U.S. (1 How.) 190, 207 (1851). This legislation impaired the obligation of contracts with respect to bank notes issued *before* the repeal, but was valid with respect to bank notes issued *after*. *See id.* at 206. ("That the state had the right to repeal the above section may be admitted. And the emissions of the bank subsequently are without the guaranty. But the notes in circulation at the time of the repeal are not affected by it."). Again, in each of these cases, statutory invalidity meant that the statute would not be given effect with respect to certain circumstances, not that the act was itself invalid. *See id.* at 207 ("The power of the legislature to repeal the section, the stock of the bank being owned by the state, is not controverted; but that act cannot affect the notes in circulation at the time of the repeal.").

111 22 U.S. (1 Wheat.) 1 (1824).

112 45 Mass. (1 Met.) 282 (1842).

113 Gibbons was the owner of two steamboats which he operated as ferries between New York City and Elizabethtown, New Jersey, pursuant to a federal license that permitted him to engage in the "coasting trade." *Gibbons*, 22 U.S. at 2. Ogden possessed the exclusive right under New York law to operate steamboat service on New York waters; he had been assigned this right by Robert Fulton and Robert Livingston, who had received it from a succession of New York statutes. *Id.* at 1–2. Ogden sued in New York state court for an injunction to enforce his exclusive right, and he obtained a permanent injunction, which he successfully defended on appeal within the New York system. *Id.* at 1–3. The Supreme Court heard Gibbons's appeal. *Id.* at 3.

Congress, and deprived a citizen of a right to which that act entitles him."[114]

In *Norris v. Boston*, Massachusetts Chief Justice Shaw combined this application-specific approach with the background principle that a statute may be valid in part and void in part. *Norris* was an assumpsit action to recover money paid under a Massachusetts law placing a per capita tax on foreign passengers brought by ship into Massachusetts ports.[115] En route to denying recovery on the ground that the statute was not repugnant to the Constitution, Chief Justice Shaw framed the applicable test (which he drew from *Gibbons*) as providing that "if the specific means [adopted by state and federal law] do not conflict with each other, both may stand and operate together, though they operate at the same time, upon the same subjects."[116] If there is any conflict, the state statute is void to the extent of the conflict but no further:

> This act having been passed in due form, and received the sanction of all branches of the legislature, . . . is valid and binding, unless it comes in conflict with the constitution of the United States or some

---

[114] *Id.* at 210 (emphasis added). This focus obviated the question of concurrent power because "[s]hould this collision exist, it will be immaterial whether [New York's] laws were passed in virtue of a concurrent power 'to regulate commerce with foreign nations and among the several States,' or in virtue of a power to regulate their domestic trade and police." *Id.* According to the Court, "[t]he real and sole question seems to be, whether a steam machine, in actual use, deprives a vessel of the privileges conferred by a license" issued under a 1793 federal statute regulating the coasting trade. *Id.* at 219. The Court concluded that "the act of a State inhibiting the use [of its waters or entering ports] to any vessel having a license under the act of Congress, comes . . . in direct collision with [the federal] act." *Id.* at 221. The Court's holding, therefore, was that "*so much of the* laws of the State of New York, *as prohibits vessels licensed according to the laws of the United States*, from navigating the waters of the State of New York, by means of fire or steam, is repugnant to the said constitution, and void." *Id.* at 240 (emphases added).

[115] 45 Mass., at 282. The Supreme Court reversed the *Norris* decision, which it decided together with *Smith v. Turner*, a similar New York case, in the *Passenger Cases*, 48 U.S. (1 How.) 283 (1849). But the Court's reasoning in reversal did not call into question Shaw's reasoning regarding partial unconstitutionality. In fact, the arguments of counsel and the opinions of the Justices in the *Passenger Cases* reveal the same provision-specific approach to partial unconstitutionality taken by Chief Justice Shaw. Section 3 of the statute at issue in *Norris* contained the invalid tax provision, while Section 1 authorized the boarding of ships for inspection, and Section 2 authorized the port city to require payment of a $1000 bond for "any lunatic, idiot, maimed, aged, or infirm person" as security against such individuals becoming "a city, town, or State charge" after landing. *Id.* at 409. The Court's constitutional reasoning invalidated only Section 3. *Id.* at 410. Indeed, both Justice Grier (in the majority) and Justice Taney (in dissent) explicitly stated that Massachusetts's power under Section 2 was not in question in the case. *See id.* at 456–57 (Grier, J.); *id.* at 469 (Taney, J., dissenting). The Court took a similarly narrow approach in *Smith v. Turner*, the companion New York case. *Id.* at 283. *Smith* was a challenge to certain sections in a New York statute requiring the payment by ship captains of a per capita tax on foreign passengers brought into the Port of the City of New York. *Id.* at 284. The Court's invalidation of the tax was limited to these sections and did not affect the other sections. *Id.* at 572.

[116] *Id.* at 294.

> law made in pursuance of its enumerated powers. *If it be in any respect repugnant thereto, to that extent it is inoperative and void*; because the constitution of the United States, and all laws and treaties made in pursuance of the powers vested in the general government, are the supreme law of the land. But *an act of the State legislature is no further void, than as its provisions are thus repugnant to the constitution or laws of the United States*; and consequently the same act may be valid in part and void in part.[117]

This approach to partial unconstitutionality limited the relevant inquiry for judicial review to the constitutionality of the specific provision giving rise to the claimed constitutional defect.[118] This limitation arose from the fact that it was only this constitutional question whose resolution was necessary to the assumpsit action in *Norris*.[119]

A second pair of cases, this time involving the Import-Export Clause, illuminates the original approach to partial unconstitutionality through another one-two combination of (i) a landmark Marshall decision on federal-state authority, followed by (ii) a later Shaw decision applying the Marshall decision. In *Brown v. Maryland*, Chief Justice Marshall formulated the "original package" doctrine to implement the Import-Export Clause.[120] This doctrine divided federal and state authority depending on whether "the importer has so acted upon the thing imported, that . . . it has perhaps, lost its distinctive character as an import, and has become subject to the taxing power of the State . . . ."[121]

---

117 *Id.* at 287–88 (emphases added).

118 *See id.* at 288. ("The question therefore in the present case, is, not whether some single provision may not be found in this act, which may conflict with the law of the United States, but whether the provision requiring the master or agent of a vessel, arriving within the limits of the State, with alien passengers on board, liable to become chargeable to the State, to pay two dollars for each of such passengers, before he can be permitted to land them, conflicts with that law.").

119 *Norris*, 45 U.S. at 284.

120 25 U.S. (1 Wheat.) 419, 441–42 (1827). In *Brown*, the Court reviewed the constitutionality of a license requirement for liquor importers and wholesalers who were prosecuted for selling without a license. *Id.* at 436. The challenged license requirement was an amendment to a statute that had previously required only retailers to pay duties. *Id.* at 450 (Thompson, J., dissenting). Chief Justice Marshall's opinion for the Court invalidated the importer/wholesaler license requirement, which functioned as a tax on imports, while leaving the retailer licensing scheme intact. *Id.* at 459. This holding depended on the Court's identification of the point at which the constitutional prohibition on state taxation of imports no longer prevented the states from levying a tax on goods that had previously been imported. Chief Justice Marshall reasoned that "when the importer has so acted upon the thing imported, that it has become incorporate and mixed up with the mass of property in the country, it has perhaps lost its distinctive character as an import, and has become subject to the taxing power of the State; but while remaining the property of the importer, in his warehouse, in the original form or package in which it was imported, a tax upon it is too plainly a duty on imports to escape the prohibition in the constitution." *Id.* at 441–42.

121 *Id.* at 441–42.

Ten years later, Chief Justice Shaw applied this doctrine in *Commonwealth v. Kimball*, in which an individual sought to overturn his conviction for unlicensed retailing of spirituous liquors.[122] Kimball argued that the Massachusetts law prohibiting such sales was an unconstitutional state regulation of interstate commerce and an unconstitutional duty on imports.[123] The court rejected both arguments.[124]

Shaw's analysis rejecting Kimball's Import-Export Clause claim provides an example of how the original approach to partial unconstitutionality accommodated laws that *by their terms* could reach more broadly than constitutionally permissible but that *in their application* were within state power. Kimball argued that the license requirement was unconstitutional because "the prohibition to sell is general, and makes no distinction between the cases of a sale by the importer of imported spirits, in the original packages . . . and the sale of spirits not imported, or not by the importer, or not in the original packages."[125] Shaw rejected this argument in two steps. First, he noted that the unconstitutionality of certain applications of the state law would have no effect on its constitutionality in other applications.[126] Second, he observed that there was no proof that the law as applied to Kimball was being used to penalize the sale of imported liquor.[127] Because

---

122 41 Mass. (1 Pick.) 359, 361–62 (1837).

123 *Id.* at 361.

124 *Id.* at 363. Addressing the Commerce Clause argument first, Shaw described the challenged law as a health and welfare regulation, which the state could permissibly enforce except to the extent that it "shall happen, in any particular instance, to come directly in conflict with the operation of some law of the United States made in pursuance of its enumerated powers." *Id.* at 361. In the case of such a conflict, the Supremacy Clause dictates that "to the extent of such collision and repugnancy, the law of the State must yield, and to that extent and no further, it is rendered by such repugnancy, inoperative and void." *Id.* Such a conflict "is not to be presumed, but . . . must be clearly shown and established." *Id.* at 365. Kimball proved no such conflict—he failed to show that the law regulating retail sales of spirituous liquors would impede federal law—and he therefore lost on this challenge. *Id.*

125 *Id.* at 362.

126 *See id.* ("Supposing the law could be construed to be repugnant to the constitution of the United States, in so far as it prohibited the sale of imported spirits by the importer in the original package, it would be void thus far and no further, and in all other respects conforming to the acknowledged power of the State government, it would be in full force."). Elaborating on this reasoning, Shaw stated further that "[w]hether legal enactments, some of which it is competent for the legislature to make, and others not, are contained in the same or in different sections of a statute, can make no difference." *Id.* The reason is that "it is not the defect of form, but of power, that invalidates any of them; it is, therefore, the subject matter, and not the arrangement of the language in which it is embodied, that is to be regarded in deciding whether any provision is constitutional or not." *Id.*

127 *Id.* ("If . . . the defendant had offered to show in his defence, that the spirits charged to have been illegally sold by [the defendant], without license, contrary to the statute, were

Kimball made no such showing, his challenge failed.[128] In other words, Shaw limited judicial review by looking to the constitutionality of state power as actually exercised under the statute, rather than by looking to the terms of the statute on their own.

The application-specific operation of judicial review in the foregoing cases depended on background principles of partial unconstitutionality. These principles contributed to a working federalism by limiting the scope of constitutional adjudication. By virtue of the displacement approach used in these cases, a state statute would remain valid and enforceable except to the extent that its application conflicted with federal law. Unlike modern decisions, as-applied adjudication did not depend on implicit or explicit satisfaction of a test for severability. Conversely, no inseverability argument was advanced to expand the scope of adjudication beyond the specific application or applications giving rise to the claim of unconstitutionality.

3. Occasions of Inseverability

A final pair of cases illustrates the distance between the earlier approach to partial unconstitutionality and the modern conception of what situations provide occasions for inseverability analysis. The contrast between an 1845 tax case and the Court's inseverability-based invalidation of the federal income tax in 1895 reveals the substantial shift in constitutional adjudication that modern severance doctrine enabled.

In the 1845 case *Gordon v. Appeal Tax Court*, the Court partially invalidated a general tax on banks that operated in some circumstances in violation of the Contracts Clause.[129] *Gordon* involved a bank tax that did not by its terms distinguish between banks whose charters exempted them from increased taxation and those whose charters did not.[130] The Court invalidated the statute as applied to banks whose charters exempted them from increased taxation while explicitly recognizing and preserving the statute's applicability to non-exempted banks.[131] Neither the Court nor any of the banks suggested

---

imported by himself, and sold in the original package, it would then have given rise to the question which has been mainly argued in the present case.").

[128] *Id.*

[129] 44 U.S. (1 How.) 133, 149–50 (1845).

[130] *Id.* at 144.

[131] *Id.* at 150; *see also* State Bank of Ohio v. Knoop, 57 U.S. (1 How.) 369 (1853) (invalidating statutory tax increase applied to banks whose charter, under prior act, provided for lower rates); Planters' Bank v. Sharp, 47 U.S. (1 How.) 301 (1848) (invalidating state restrictions on negotiability of bank notes because prior act required that they be negotiable). Because the doctrinal test for invalidity under the Contracts Clause at the time looked only at specific contracts at issue, the extent of invalidity in *Knoop* was limited to

that the unconstitutionality of certain applications of the law could also result in the invalidity of the law in its constitutional applications through the use of inseverability reasoning.

The absence of such inseverability arguments from *Gordon* is uninteresting given the then-prevailing approach to partial unconstitutionality, but it is striking in comparison with the later decision in *Pollock v. Farmer's Loan & Trust Co*.[132] In this 1895 decision, the Supreme Court voted 5-4 to invalidate entirely the income tax portions of an 1894 revenue act.[133] The Court first held that the tax was unconstitutional insofar as it was based on income from both real and personal property.[134] The Court next concluded that the remainder of the income tax was inseparable from the tax on income derived from real and personal property; thus, the income tax had to be struck down in its entirety.[135] Writing for the Court, Chief Justice Fuller reasoned that the holding of unconstitutionality regarding income derived from real and personal property would eliminate "by far the largest part of the anticipated revenue" from the income tax, leaving "the burden of the tax to be borne by professions, trades, employments, or vocations . . . ."[136] Eliminating only the unconstitutional portions of the income tax, then, would cause "what was intended as a tax on capital" to become instead "a tax on occupations and labor."[137]

---

the statute's applications to only those contracts actually impaired: "[T]he tax law of 1851 . . . impairs the obligation of the contract, which is prohibited by the Constitution of the United States, and, consequently, that *the act of 1851, as regards the tax thus imposed, is void*." *Id.* at 392 (emphasis added). Similar reasoning appears in *Planters' Bank*. *See* 47 U.S. at 334 ("[T]he law under which this action has been abated must be considered as having impaired the obligation of contracts, and therefore to be *in this respect* unconstitutional, and the judgment of the State court erroneous.") (emphasis added).

132 158 U.S. 601 (1895).

133 The Court heard argument twice in *Pollock*. Its first decision did not resolve the entire case because the Court was evenly divided on several questions. *See* Pollock v. Farmers' Loan & Trust Co., 157 U.S. 429, 586 (1895). In its first decision, the Court held that a tax on rents or incomes from real estate was unconstitutional. *Id.* But the Court was evenly divided on the remaining questions, which included: (1) whether the invalidity of the tax on rents and income from real property invalidated the whole act; (2) whether the tax on income derived from personal property was unconstitutional; and (3) whether any other parts of the income tax were unconstitutional. *Id.* Because the effect of an even division is to affirm the judgment below, which in this case had rejected the constitutional challenge in its entirety, the only portion of the law that was unconstitutional after the first decision was the income tax on rents and income from real estate. *Id.* The second decision is discussed in the text above.

134 *Pollock*, 158 U.S. at 635.

135 *Id.* at 637.

136 *Id.*

137 *Id.*

Because such a tax was not intended by Congress, the Court held that the income tax provisions had to fall in their entirety.[138]

Discussing *Pollock* a few decades later, Robert Stern observed the close connection between the Justices' severability analyses and their views on the merits of the case: "The judges who disapproved of the law thought it was inseparable, while those who regarded it as constitutional also took a contrary position on the question of separability."[139] By that time, Stern was working within the legislative-intent approach to severability, however, so he did not appreciate just how far the Court's reasoning departed from its earlier approach to partial unconstitutionality. Under that earlier approach, the idea that the resulting relative tax burden from a holding of partial unconstitutionality should be a factor in the Court's consideration of the validity of the income tax as a whole was, quite literally, inconceivable.

C. The Stealthy Rise of Modern Severability Doctrine

The foregoing cases reveal an internally coherent, transsubstantive approach to the problem of partial unconstitutionality. The cases are not idiosyncratic but instead are part of a general framework applied in both state and federal courts for several decades prior to the rise of severability doctrine. A basic principle governed: Statutes are invalid so far as they are repugnant to superior law, but no further.[140]

Why has this original approach to partial unconstitutionality languished in obscurity? One reason is framing. Rather than view *partial*

---

[138] *Id.*

[139] Stern, *supra* note 15, at 113.

[140] Although I have not undertaken to research the approach to partially unconstitutional (or otherwise partially invalid) statutes before *Marbury*, the limited evidence that I have come across is consistent with the approach described above. *See* Hutchins v. Player, (1663) 124 Eng. Rep. 585, 610 (C.P.) ("[T]hat part of the act which concerns pitching cloths at Blackwell Hall, and the hallage, is purely distinct, and hath no relation to that part of the act which concerns the factors; and, therefore, the one part, as I conceive, shall not vitiate the other. And this construction of the act, as it is agreeable with common reason, that what is good shall not be vitiated by what is bad, so it is also agreeable with the reason of the law."); JOSEPH HENRY SMITH, APPEALS TO THE PRIVY COUNCIL FROM THE AMERICAN PLANTATIONS 628 (1950) (reporting on 1760 opinion by two crown law officers that "the King in Council lacked the power to disallow in part Pennsylvania acts," but that "there may be cases in which particular provisions may be void ab initio though other parts of the law may be valid, as in clauses where any act of Parliament may be contraversed or any legal right of a private subject bound without his consent"); 1 JOHN STRANGE, REPORTS OF ADJUDGED CASES IN THE COURTS OF CHANCERY, KING'S BENCH, COMMON PLEAS, AND EXCHEQUER 469 (London, A. Strahan & W. Woodfall 1795) ("There is no doubt but a by-law may be good in part, and void for the rest; for where it consists of several particulars, it is to all purposes as several by-laws, though the provisions are thrown together under the form of one." (reporting on Fazakerley v. Wiltshire, (1716) 88 Eng. Rep. 754 (K.B.))); *see also* HAMBURGER, LAW AND JUDICIAL DUTY 262 n.11 (citing cases).

unconstitutionality as a phenomenon in its own right, Stern's foundational scholarship took *severability doctrine* as a given, traced it back to its origins, and dismissed the earlier framework by observing that "[t]he problem of separability, as we know it today, was not recognized" in those earlier years.[141] That approach made sense for Stern; he sought to bring order to the modern doctrine, not discard it.

The problem from this framing comes in the shift from Stern's position[142] to the more recent claim that "[s]everability doctrine has been with us since the beginnings of judicial review."[143] This anachronistic attribution of implicit severability all the way back to *Marbury* accommodates that case within the now-prevailing excision-based approach, but it obscures the original displacement-based approach that *Marbury* actually exemplifies.

Another reason for the obscurity of the original approach is that the background principles informing the original approach to partial unconstitutionality were so uncontroversial that they did their work without much discussion in the Court's opinions. Severability doctrine now has a name and can easily be identified. By contrast, the background principles regarding partial unconstitutionality in the first several decades of judicial review were not grouped together in their own named doctrinal category. Nobody argued against these principles. As a result, the identification of partial unconstitutionality as a problem for analysis in its own right did not occur until after the modern intent-based approach to severability emerged.

In 1854, Chief Justice Lemuel Shaw of the Massachusetts Supreme Judicial Court delivered an opinion in *Warren v. Mayor & Aldermen of Charlestown*[144] that Stern credits as the first decision to hold, on the basis of legislative intent, that the invalid parts of a statute could nullify the statute in its entirety.[145] This doctrinal innova-

---

[141] Stern, *supra* note 15, at 79. Stern's use of "separability" rather than "severability" reflected the more prevalent usage when Stern wrote in 1937. The terms are understood today to be interchangeable. *See* Fallon, *supra* note 6, at 1331 n.55 ("The terms 'severable' and 'separable,' and 'severability' and 'separability,' are treated as synonymous by the courts and in the literature . . . .").

[142] Stern, *supra* note 15, at 79.

[143] Gans, *supra* note 55, at 639; *see also* Dorf, *supra* note 26, at 250 (describing Court's opinion in *Marbury* as applying "the same presumption of severability" as in *Yazoo*); Shumsky, *supra* note 15, at 232 (describing *Marbury* as reflecting "early assumption that partially unconstitutional statutes were to be severed").

[144] 68 Mass. (2 Gray) 84, 94 (1854).

[145] Stern, *supra* note 15, at 79–80. It turns out that *Warren* was not, as Stern thought, the first decision to use a legislative intent test to hold "that the invalid parts of a law might nullify the remainder." *Id.* Others have followed Stern's error on this point. *See, e.g.*, Nagle, *supra* note 15, at 212; Shumsky, *supra* note 15, at 233. In actuality, the Arkansas Supreme Court decided two cases before *Warren*, one in 1851 and another in 1853, in which the court relied on legislative intent to conclude that statutes' unconstitutional parts

tion was to provide a way for a court to hold entirely invalid a statute that was only partially unconstitutional, based on judicial perception of whether the legislature would have enacted the statute absent its unconstitutional parts. Stern summarized the legislative-intent test of *Warren* as follows:

> [I]f constitutional and unconstitutional portions of a statute [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' the entire statute must fall.[146]

Stern's summary accurately captures the manner in which *Warren* rendered severability dependent on legislative intent. But Stern did not explain from where Shaw drew this test. A clue comes from the omission in Stern's quotation: The omitted language in the quote required the reviewing court to determine whether the unconstitutional and constitutional parts of the challenged statute were dependent on each other "as conditions, considerations or compensations for each other."[147] As Mark Movsesian has explained, "Chief Justice Shaw's reference in *Warren* to 'conditions, considerations, or compensations' is significant. These, of course, are basic concepts of contracts law."[148] In *Warren*, then, Shaw rendered statutory severability dependent on a legislative intent–based test drawn from contract law, apparently based on the implicit assumption that a statute is "a bargain among the legislators who enacted it."[149] The limiting principle, in Movsesian's words, is that a "court could not sever a provision so central as to amount to 'consideration' for the statute's passage."[150] Shaw thus advanced a new approach to partial unconstitutionality through his *Warren* decision.

Although Shaw's development of a contract law–inspired, intent-based test for severability marked a shift from the earlier approach to partial unconstitutionality, the shift was subtle. The legislative-intent limitation attributed to *Warren* did not spring fully formed from Chief

---

were inseparable from their constitutional parts, requiring invalidation of the entire statutes. *See* Washington v. State, 13 Ark. 752, 763–64 (1853) (invalidating billiards licensing scheme); Eason v. State, 11 Ark. 481, 501–03 (1851) (invalidating act expanding jurisdiction of justices of peace). But there is no point in getting hung up on figuring out precisely what was first in time: The Arkansas cases proved to be less influential than *Warren*, which appears first in eminence if not in time; and any other earlier cases had to have been even less influential than the Arkansas cases.

146 Stern, *supra* note 15, at 80 (quoting *Warren*, 68 Mass. (2 Gray) at 99).

147 *Warren*, 68 Mass. (2 Gray) at 99.

148 Movsesian, *supra* note 15, at 62.

149 *Id.*

150 *Id.*

Justice Shaw's mind but was prefigured by earlier formulations that were susceptible to transmutation into the legislative-intent approach. These formulations appeared in cases that applied a "void in part, valid in remainder" rule of decision, but did so in a way that presaged the future legislative-intent limitation.

Clark v. Ellis, an 1826 decision of the Indiana Supreme Court, supplies one such formulation.[151] The issue in *Clark* was whether an assault-and-battery prosecution was within the jurisdiction of the justice of the peace who had held the trial and delivered the sentence.[152] An Indiana statute vested justices of the peace with jurisdiction to preside over jury trials for assaults and batteries in cases resulting in fines of no more than twenty dollars.[153] But the Indiana Constitution limited the justice of the peace's jurisdiction to certain cases resulting in fines of no more than three dollars.[154] The court held that the justice of the peace acted within his constitutional jurisdiction because the decision giving rise to the constitutional challenge had resulted in a fine of three dollars.[155] In explaining why the exercise of jurisdiction in that case was constitutional even though the terms of the statutory grant exceeded the constitutional limitation, the court stated that "a part of an act of assembly being unconstitutional, does not affect a constitutional part of the same act relative to the same subject."[156] Had the court stopped with this statement, there would have been little in *Clark* for transmutation into modern severability doctrine. But the court continued: "That part which is unconstitutional, is considered as if stricken out of the act; and if enough remains to be intelligibly acted upon, it is considered as the law of the land."[157] It is but a short step from this idea that the unconstitutional part "is considered as if stricken out of the act" to the modern notion that judges should look to legislative intent to determine whether to excise provisions from statutes.

Another early state case, *Campbell v. Mississippi Union Bank*, also illustrates how the original displacement approach was suscep-

---

151 2 Blackf. 8 (Ind. 1826).

152 *Id.* at 9. *Clark* was not itself the prosecution but apparently a defamation action that followed that prosecution. *Id.* at 8–9. The plaintiff, Clark, had testified against Ellis before the justice of the peace, and Ellis accused Clark of committing perjury. *Id.* at 9. Clark then sued Ellis, and one of the elements of Clark's cause of action required showing that Ellis's accusations "were spoken with reference to a swearing" in the trial before the justice of the peace. *Id.* Ellis demurred on the ground that the proceeding before the justice of the peace was outside of the latter's jurisdiction. *Id.*

153 *Id.*

154 *Id.*

155 *Id.* at 9–10.

156 *Id.* at 10.

157 *Id.*

tible to transmutation into the modern intent-based excision approach.[158] In discussing the relationship between two acts passed at different times (the latter of which was supplemental to the first one), the court explained that a previous valid act could not be destroyed by a subsequent void one.[159] Said the court: "If one be void the other will not be affected, unless they be so essentially connected and blended together as to make one useless and inoperative without the other."[160] By introducing a qualification to the general principle that "[i]f one be void the other will not be affected," this decision presages modern severability doctrine's intent-based test. Admittedly, modern severability doctrine goes beyond asking whether what remains would be "useless and inoperative" to inquiring into the hypothetical intent of the enacting legislature. But the fact that *Campbell* conceptualizes the "useless and inoperative" remainder as a qualification of the general principle of void only so far as repugnant—rather than as an independent reason why the remainder could not be enforced as law—reveals the short conceptual distance between the earlier approach to partial constitutionality and modern severability doctrine.

In any event, the novelty of Chief Justice Shaw's invocation of legislative intent as a test for severability was not recognized at the time it was introduced, or even in the immediately following decades. Thomas Cooley, in the 1868 first edition of his influential treatise, matter-of-factly included a legislative intent–based test for severability in his discussion of partial unconstitutionality, supported with an impressive citation of approximately two dozen cases (only some of which actually used the legislative-intent approach).[161] Cooley's recitation of the legislative intent approach portrayed it as seamlessly continuous with what preceded it, which effectively buried the earlier approach to partial unconstitutionality.

A comparison of the first (1857)[162] and second (1874)[163] editions of Theodore Sedgwick's *A Treatise on the Rules Which Govern the*

---

158 7 Miss. (6 Howard) 625 (Err. & App. 1842).

159 *Id.* at 676–77.

160 *Id.* at 677.

161 THOMAS M. COOLEY, A TREATISE ON THE CONSTITUTIONAL LIMITATIONS WHICH REST UPON THE LEGISLATIVE POWER OF THE STATES OF THE AMERICAN UNION 177–81 (1st ed. Boston, Little, Brown 1868).

162 THEODORE SEDGWICK, A TREATISE ON THE RULES WHICH GOVERN THE INTERPRETATION AND APPLICATION OF STATUTORY AND CONSTITUTIONAL LAW (1st ed. New York, Voorhies 1857). Born to a prominent American family in 1811, Sedgwick was a New York lawyer who authored a treatise on damages in addition to this treatise on interpretation. THE YALE BIOGRAPHICAL DICTIONARY OF AMERICAN LAW 487–88 (Roger K. Newman, ed., 2009).

163 THEODORE SEDGWICK, A TREATISE ON THE RULES WHICH GOVERN THE INTERPRETATION AND APPLICATION OF STATUTORY AND CONSTITUTIONAL LAW (John Norton Pom-

Interpretation and Application of Statutory and Constitutional Law reveals the quickness and stealth of the incorporation of the legislative intent test into constitutional adjudication.

The first edition of Sedgwick's treatise states: "The principle that a statute is void only so far as its provisions are repugnant to the constitution, that one provision may thus be void and this not affect other provisions of the statute, has been frequently declared."[164] In addition, Sedgwick block-quotes an 1854 decision of the Massachusetts Supreme Judicial Court authored by Chief Justice Lemuel Shaw:

> 'The principle is now well understood,' says the Supreme Court of the State of Massachusetts, 'that where a statute has been passed by the legislature under all the forms and sanctions requisite to the making of laws, some part of which is not within the competency of the legislative power, or is repugnant to any provision of the constitution, such part thereof will be adjudged void and of no avail; whilst all other parts of the act, not obnoxious to the same objection, will be held valid and have the force of law. There is nothing inconsistent in declaring one part of the same statute valid and another part void.'[165]

That is the entirety of Sedgwick's discussion: one sentence stating a general principle followed by a block quote from Chief Justice Shaw. There is no mention of "separability," "severability," or "severance," nor qualifications of this general principle.

Although Sedgwick's description is spare, three features are worth noting. First, the only function of the principle that "void in one part does not mean void in any other parts" is to limit the scope of judicial invalidation. Because of this principle, the taint of unconstitu-

---

eroy ed., 2d ed. New York, Baker, Voorhies 1874). Pomeroy was a professor at the New York University School of Law and the Hastings School of Law, as well as a prolific treatise writer, editor, and case reporter, in the last third of the nineteenth century. *See* John C. Leary, *John Norton Pomeroy 1828–1885: A Biographical Sketch*, 47 L. LIBR. J. 138, 139–42 (1954).

[164] SEDGWICK, *supra* note 162, at 489. Sedgwick's authorities were: Ely v. Thompson, 10 Ky. (3 A.K. Marsh) 70 (1820); Gibbons v. Ogden, 22 U.S. (9 Wheat.) 1 (1824); Clark v. Ellis, 2 Blackf. 8 (Ind. 1826); City of New York v. Miln, 36 U.S. (11 Pet.) 102 (1837); Commonwealth v. Kimball, 41 Mass. (24 Pick.) 359 (1837); Edwards v. Pope, 4 Ill. (3 Scam.) 465 (1842); Norris v. Boston, 45 Mass. (4 Met.) 282 (1842); and Fisher v. McGirr, 67 Mass. (1 Gray) 1 (1854). But the profusion of case citations can be misleading. Sedgwick's cases are spread across two footnotes. For each footnote, the lead cited case contains the principle articulated by Sedgwick and then itself cites the other cases that appear in Sedgwick's footnotes. In effect, then, Sedgwick relied on two cases, *Edwards* for the first sentence and *Fisher* for the block quote. *Id.*

[165] SEDGWICK, *supra* note 162, at 489 (quoting *Fisher*, 67 Mass. (1 Gray) at 21–22). Although decided in the same year as *Warren*, *Fisher* preceded that watershed case. *Fisher* was decided in the Supreme Judicial Court's March 1854 term, while *Warren* was decided in the October 1854 term. *See Fisher*, 67 Mass. (1 Gray) at 1; *Warren*, 68 Mass. (2 Gray) at 84.

tionality and consequent voidness attach only to the extent that the statute's individual terms and applications are unconstitutional. Second, this principle applies to a statute that "has been passed by the legislature under all the forms and sanctions requisite to the making of laws."[166] Its application presupposes a statute that is properly in existence as law. The principle provides for the continued existence of some parts or provisions notwithstanding that other parts or provisions are void. Third, the principle does not depend on legislative intent.

Sedgwick's 1874 second edition retains the same text as the first edition but expands its discussion with an extensive note that discusses the "question" of separability and the doctrinal test for determining whether a statute is separable.[167] In less than two decades, then, a brightline principle that an unconstitutional statutory provision has no effect on other constitutional provisions in the same statute was transformed into a problem that only posed the "question" of separability to be resolved by application of a doctrinal test. The shift to legislative intent–based severability doctrine had already taken place.

The Supreme Court handed down the next landmark decision for modern severability doctrine, *United States v. Reese*,[168] the following year. The doctrinal innovation in *Reese* was the Supreme Court's embrace of the understanding that, by enjoining only the unconstitutional applications of a statute that has valid and invalid applications, a court may engage in forbidden judicial rewriting.[169] *Reese* involved the federal prosecution of two municipal election inspectors in Kentucky for refusing to count the vote of William Garner, an African American.[170] The Supreme Court upheld dismissal of the indictment on the ground that the statute the officials were charged with violating was not appropriate enforcement legislation under the Fifteenth Amendment.[171] The Court recognized Congress's power to enact legislation punishing interferences with voting by state officials on account of race but interpreted the statute to punish *all* interferences with voting, regardless of whether the interferences were on account of race.[172] As thus interpreted, the statute was unconstitutional. It therefore could not form the basis of the indictments in *Reese*, even

166 SEDGWICK, *supra* note 162, at 489 (quoting *Fisher*, 67 Mass. (1 Gray) at 21).

167 SEDGWICK, *supra* note 163, at 413–15, 413–14 n.(a).

168 92 U.S. 214 (1875).

169 *See id.* at 221 ("The question, then, to be determined, is, whether we can introduce words of limitation into a penal statute so as to make it specific, when, as expressed, it is general only.").

170 *Id.* at 215.

171 *Id.* at 221–22.

172 *Id.* at 218, 220.

though those indictments charged the denial of an individual's right to vote on account of his race.[173] The Court stated, "[w]e are not able to reject a part which is unconstitutional, and retain the remainder, because it is not possible to separate that which is unconstitutional, if there be any such, from that which is not."[174] This conclusion followed from the Court's earlier determination that "[i]f, taking the whole statute together, it is apparent that it was not the intention of Congress thus to limit the operation of the act, we cannot give it that effect."[175]

The link in *Reese* between this intent-focused approach to statutory interpretation and the Court's conclusion as to inseparability was the Court's conception of judicial review. The Court reasoned that the "[t]he proposed effect is not to be attained by striking out or disregarding words . . . , but by inserting those that are not now there."[176] Having framed the question as "whether [the Court] can introduce words of limitation into a penal statute so as to make it specific, when, as expressed, it is general only," the Court answered: "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."[177] In short, the Court declined to apply severability reasoning on an application-by-application basis here, stating that confining the statute to only its constitutional applications would have required the Court to write in words of limitation, which the Court took to be beyond the judicial power.

The importance of *Reese* for our purposes lies not in its substantive constitutional holding but in the fact that it equates the idea of giving effect to only the constitutional applications of a partially unconstitutional statute with "mak[ing] a new law."[178] The Court has continued to rely on *Reese*'s caution against "substitut[ing] the judicial for the legislative department of the government"[179] through trimming of a partially unconstitutional statute to leave only its constitutional applications.[180] In so doing, however, the Court has not

---

[173] *Id.* at 215.
[174] *Reese*, 92 U.S. at 221.
[175] *Id.* at 219.
[176] *Id.* at 221.
[177] *Id.*
[178] *Id.*
[179] *Id.*
[180] *See* Ayotte v. Planned Parenthood of N. New England, 546 U.S. 320, 330 (2006) ("[The Court is] wary of legislatures who would rely on our intervention . . . ."); City of Chicago v. Morales, 527 U.S. 41, 60 (1999) ("The Constitution does not permit a legislature to 'set a net large enough to catch all possible offenders, and leave it to the courts to step inside and say who could be rightfully detained, and who should be set at large.'" (quoting *Reese*, 42 U.S. at 221)); Reno v. ACLU, 521 U.S. 844, 884–85 n.49 (1997) (same).

reconciled that reliance with its 1960 decision in *United States v. Raines*, which adopted an approach to the constitutionality of enforcement legislation opposite that of *Reese*.[181]

In present times, the excision-based legislative-intent approach is so pervasive that West's Key system describes that approach as present in the early nineteenth century even when it was not.[182] And while the coverage of severability doctrine in treatises is heavy with case law, it is also uneven. For example, the 2001 sixth edition of Sutherland's statutory interpretation treatise continues to describe *Reese* as a "leading federal case"[183] without any mention of the Court's rejection of *Reese*'s approach to the constitutionality of federal enforcement legislation in *Raines*. The fact is that it has become difficult to conceive of judicial review in any way other than excision.

* * *

Because the excision-based legislative-intent approach is now so ingrained, it might be tempting to minimize the difference between that approach and the original displacement approach by describing the original approach in excision terms—for instance, by asserting that the original approach represents nothing more than a blanket rule of severability. This temptation should be avoided. Ignoring the distinctive displacement-based framing of the original approach to partial unconstitutionality obscures the significant change in the conception of the judicial role that the transition from displacement to excision represents. Moreover, neglect of the displacement approach as an alternative to excision has resulted in a failure of scholars and judges to apprehend the distinctively destructive function of modern severance doctrine.

This Article's recovery of the original displacement approach to partial unconstitutionality enables a before-and-after comparison that upends the dominant view that severability doctrine saves partially

---

[181] 362 U.S. 17, 24 (1960) ("[T]o the extent *Reese* did depend on an approach inconsistent with what we think the better one and the one established by the weightiest of the subsequent cases, we cannot follow it here."); *id.* at 24–25 ("[I]f the complaint here called for an application of the statute clearly constitutional under the Fifteenth Amendment, that should have been an end to the question of constitutionality.").

[182] For example, the key-cited headnote for an 1826 state court decision that does not so much as mention legislative intent describes the case as if the reasoning and disposition on severability turned on it. *See* Clark v. Ellis, 2 Blackf. 8, 1826 WL 1078 (Ind. 1826) ("Statute may be valid in part and invalid in part, and invalid part may be disregarded, where the two parts are not so intimately connected as to raise presumption that Legislature would not have enacted one without other.").

[183] 2 NORMAN J. SINGER, SUTHERLAND STATUTES AND STATUTORY CONSTRUCTION 608 (6th ed. 2001).

unconstitutional statutes from total invalidity.[184] *Before* severance doctrine emerged, displacement-based partial unconstitutionality doctrine was already doing precisely this. Prior to severability doctrine's rise, courts routinely held an unconstitutional law void to the extent of repugnancy, but no further; there was no "next step" in which courts inquired into whether the legislature would have preferred no law at all to the constitutional remainder. *After* severance emerged, partial unconstitutionality (now seen as resulting from excision, and not displacement) became dependent on the satisfaction of a legislative-intent test. The "next step" of a legislative-intent analysis injects into judicial review the possibility that invalidity extends beyond unconstitutionality, which occurs whenever severability doctrine results in a holding of inseverability. In comparison with the original displacement approach, the distinctive function of modern severability doctrine is to enable *total* invalidity to flow from *partial* unconstitutionality. Modern severability doctrine does not save; it destroys.

III
DISPLACEMENT

Given that severability is not inevitable, and that existing doctrine is destructive and manipulable, the question arises: Is there a better way? There is, and it is easy to implement. This Part describes the approach of "displacement without inferred fallback law," explains why it is superior, identifies some of its limitations, and answers potential objections.[185]

---

[184] *See* Dorf, *supra* note 8, at 370 (arguing that, without severability, "any judicial decision finding any law unconstitutional, on its face or as applied, would call into question the entire legal code"); Metzger, *supra* note 18, at 887–88 ("If unconstitutional applications are not severed, the statute cannot be applied to any litigant, even one making no claim of constitutional protection for her conduct.").

[185] This Article does not advocate any particular theory of constitutional interpretation, such as originalism, nor does it argue that its proposed approach—displacement without inferred fallback law—should be adopted because it is more faithful to the original approach to partial unconstitutionality. The object of this Article is to retrieve a "useable past" through the study of a neglected aspect of our constitutional tradition. *See* Cass R. Sunstein, *The Idea of a Useable Past*, 95 COLUM. L. REV. 601, 603 (1995) (describing idea of "useable past" as oriented to "the goal of finding elements in history that can be brought fruitfully to bear on current problems," and defending constitutional lawyers' search for a useable past as distinct from "mere advocacy" or "history lite"). The method is not simply to identify a discrete set of legal directives that can be ripped from the world of early- to mid-nineteenth-century constitutional adjudication and burned into existing doctrine. Instead, the method begins by trying to understand and appreciate how courts during that time dealt with a practical problem that remains today, although understood within a different conceptual framework. It continues by correlating the practical solution from that time period with the conceptual framework and doctrinal tools available in present times.

A. Displacement Without Inferred Fallback Law

The alternative approach to partial unconstitutionality that I propose is displacement without inferred fallback law. This proposal (i) replaces the excision-based approach to judicial review with a displacement-based approach, and (ii) limits the scope of displacement by prohibiting inferences about counterfactual legislative intent used to invalidate beyond the required displacement. Under this displacement-based approach, judicial review is an exercise in determining the extent to which superior law displaces inferior law. Judicial review calls for an "exercise of judicial discretion, in determining between two contradictory laws."[186] Within this displacement approach, the reviewing court does nothing to the inferior law except refuse to recognize it as enforceable law in resolving the particular case before it.

A helpful way to think about the move from excision to displacement is as the shift from a *statute-minus* approach to a *statute-plus* approach. Laurence Tribe described the contrast between these two approaches in his analysis of the Court's invalidation and severance of the legislative veto in *Chadha*:[187]

> Rather than conceiving of the court as enforcing the law 'minus' its invalidated provision—a law the legislature never enacted—perhaps one should simply understand the court as resolving the controversy before it in terms of the *entire* body of law applicable to that controversy, the entire Act of Congress (*not* the Act 'minus' any offending portion) *plus the Constitution*.[188]

Tribe describes this approach in passing as "the theory of *Marbury v. Madison*,"[189] but does not elaborate further. Moreover, I have found no scholarship or modern judicial opinion that has explicitly developed the statute-plus approach beyond Tribe's identification of it in his analysis of *Chadha*. As the elaboration of the original approach to partial unconstitutionality in Part II has revealed, how-

---

Additionally, although the proposed approach to partial unconstitutionality includes a textualist approach to statutory interpretation, its adoption does not require the adoption of textualism more generally. Even nontextualists may conclude that the use of imaginary reconstruction for deciding what to do with partially unconstitutional statutes should be abandoned for reasons specific to the particular context of discerning hypothetical intent regarding partial unconstitutionality. As matters now stand, however, the disjunction between textualism and severability doctrine runs in the other direction. *See* Caleb E. Nelson, *What Is Textualism?*, 91 VA. L. REV. 347, 404–05 (2005) (explaining that "[t]extualist judges regularly join opinions" that use "a species of imaginative reconstruction" to answer questions of severability).

186 THE FEDERALIST No. 78, at 380 (Alexander Hamilton) (Terence Bell ed., 2003).

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

188 LAURENCE H. TRIBE, CONSTITUTIONAL CHOICES 81–82 (1985).

189 *Id.* at 318 n.116.

ever, this statute-plus approach does indeed reflect "the theory of *Marbury v. Madison*."[190]

Under this proposed displacement approach, the combination of four basic background principles drawn from the original approach to partial unconstitutionality will generally cause the scope of displacement to be coextensive with the scope of unconstitutionality. First, the function of judicial review should be to determine what law governs in the course of resolving a particular case. More specifically, judicial review should be conceived of as the process of determining whether, and, if so, to what extent, the Constitution supersedes otherwise operative law in resolving the particular case before the court.[191] Second, upon recognition of a conflict, superior law must displace inferior law, which is thereby deemed void.[192] Third, a law may be void in part and valid in part, and the voidness of one part has no effect on the validity of the other parts.[193] Fourth, a law is void to the extent of its conflict with the Constitution, but no further.[194]

Nothing in these principles rules out the possibility of facial invalidation flowing from application of substantive constitutional law. The application of some constitutional tests results in a holding that the challenged statute is completely invalid.[195] Recall, for instance, Justice Bushrod Washington's holding in *Golden v. Prince* that a state debtor-relief law was "altogether unconstitutional" because the power to

---

190 *Id.*

191 *See, e.g.*, Golden v. Prince, 10 F. Cas. 542, 544 (C.C.D. Pa. 1814) (No. 5509) (framing issue as whether contested law "is repugnant to the constitution . . . and, on that account, is not to be regarded by the court, in this case").

192 *See, e.g.*, Marbury v. Madison, 5 U.S. (1 Cranch) 137, 178 (1803) (describing resolution of conflicting law as "very essence of judicial duty").

193 *E.g.*, SEDGWICK, *supra* note 162, at 489.

194 *E.g.*, Norris v. Boston, 45 Mass. (4 Met.) 282, 287–88 (1842). A narrow qualifier on this principle is that the remainder will not survive as enforceable law if it is incapable of functioning independently of the unconstitutional provision or provisions of a statute. This qualifier states something of a truism, but is worth noting nonetheless. Its presence in case law predates the advent of the legislative intent–based approach to severability. *See, e.g.*, Clark v. Ellis, 2 Blackf. 8, 10 (Ind. 1826) ("That part which is unconstitutional, is considered as if stricken out of the act; and if enough remains to be intelligibly acted upon, it is considered as the law of the land."). The same qualifier is present in current law, although it is now understood in legislative intent–based terms. *See, e.g.*, Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 684 (1987) ("Congress could not have intended a constitutionally flawed provision to be severed from the remainder of the statute if the balance of the legislation is incapable of functioning independently."). This qualifier states a very narrow limitation on the principle of "void to the extent of repugnancy, but no further." The test is not whether the remainder would function *in the manner intended by the legislature*, but whether the remainder is capable of functioning at all.

195 *See, e.g.*, Fallon, *supra* note 6, at 1338 (describing doctrinal tests that result in facial invalidation, such as "purpose" tests and "suspect-content" tests).

make bankruptcy laws was exclusively federal.[196] In such a circumstance, there is no resulting partial unconstitutionality and no need for a court to adopt any particular approach to partially unconstitutional statutes.[197]

Just as there is nothing that forecloses facial invalidation as a matter of constitutional law, there is nothing in the displacement approach that would foreclose the legislature from enacting, in advance of invalidation, fallback law that dictates what effect partial unconstitutionality will have on the remainder of the statute. For example, the legislature could link provisions with an inseverability clause, so that they stand or fall together.[198] Absent some constitutional defect with the inseverability clause itself, the two linked provisions would live or die as a single unit even under the displacement approach with no inferred fallback law. Alternatively, the legislature could explicitly provide for some other type of fallback law, such as a provision to be substituted in place of one found to be invalid.[199] When the legislature actually provides for such fallback law, the scope of displacement will not be coextensive with the scope of unconstitutionality but will instead depend on what the legislature has actually provided.

---

196 *Golden*, 10 F. Cas. at 547.

197 This is actually not always true. To be more precise, even when a challenged statute is held facially invalid because the applicable test of substantive constitutional law requires such a holding, it may be necessary to invoke some approach to partial unconstitutionality to limit the spillover effects of the constitutional holding. For example, the Child Online Protection Act held facially unconstitutional under the First Amendment in *ACLU v. Mukasey*, 534 F.3d 181, 207 (3d Cir. 2008), was not a stand-alone statute but Title XIV of an omnibus appropriations bill. *See* Pub. L. No. 105-277, 112 Stat. 2681 (1998). The First Amendment defects that required facial invalidation of the Child Online Protection Act had no effect on the other fifty-two titles in the omnibus bill.

198 *See generally* Friedman, *supra* note 23, at 909–17 (examining distinctive features of inseverability clauses). Unlike an inseverability clause or a substitute provision, a severability clause would yield the same result as the *absence* of any fallback law. For that reason, severability clauses are superfluous under my proposed approach. However, severability clauses will not be superfluous as long as courts addressing questions of partial unconstitutionality continue to adopt the current approach of imaginative reconstruction for severability questions. As long as *inferences* of inseverability or of some other type of fallback law are permitted, severability clauses can serve to forestall those inferences and are therefore not superfluous.

199 For example, the Bipartisan Campaign Finance Reform Act contained a "'backup' definition of 'electioneering communication,' which would become effective if the primary definition were 'held to be constitutionally insufficient by final judicial decision to support the regulation provided herein.'" McConnell v. FEC, 540 U.S. 93, 190 n.73 (quoting 2 U.S.C. § 434(f)(3)(A)(ii)). Similarly, the statute at issue in Bowsher v. Synar, 478 U.S. 714 (1986), contained a "fallback" provision to take effect if a certain statutory provision were held invalid. *Id.* at 735 (holding that "fallback" provisions in Balanced Budget and Emergency Deficit Control Act were effective after statute was held unconstitutional).

Because a displacement approach allows the legislature to provide for fallback law, it must be linked with an approach to statutory interpretation for determining *whether* the legislature has in fact provided such fallback law. When looking for fallback law, the limit imposed by my proposed approach is that a court should not ask what the legislature *would have done* with respect to fallback law, only what it *actually did* with respect to fallback law. In other words, the court must be persuaded that the legislature *actually legislated a fallback provision* before engaging in anything other than simple displacement.

For several reasons discussed in the next section, this Article proposes a clear statement requirement for legislating fallback law. Before turning to that discussion, however, this exposition of how the proposed approach would operate concludes with an example that illustrates how displacement without inferred fallback law would have required a different outcome than existing severability doctrine in a recent, high-profile decision. In addition to explaining how my proposed approach would function, this example also suggests some of the advantages and disadvantages of displacement without inferred fallback law.

Bismullah v. Gates[200] was a proceeding brought by detainees seeking release from Guantánamo Bay pursuant to the procedure set forth in the Detainee Treatment Act (DTA)—a federal statute enacted in the wake of the Supreme Court's holding in *Rasul v. Bush*[201] that statutory jurisdiction to issue writs of habeas corpus extended to Guantánamo Bay. The DTA eliminated habeas review for noncitizens held at Guantánamo Bay[202] and put in its place a system of military review by Combatant Status Review Tribunals (CSRTs) followed by review in the D.C. Circuit.[203]

In the Supreme Court's first encounter with the DTA, in *Hamdan v. Rumsfeld*,[204] it held that Congress did not intend the elimination of habeas corpus review to apply retroactively to cases that had been filed prior to the enactment of the DTA. Congress responded to *Hamdan* with the Military Commissions Act (MCA), which eliminated habeas corpus review again and made clear that this elimination extended to all habeas corpus petitions on behalf of noncitizens held

---

200 551 F.3d 1068 (D.C. Cir. 2009).

201 542 U.S. 466, 484 (2004).

202 Detainee Treatment Act of 2005, Pub. L. No. 109-163, § 1405(e)(1), 119 Stat. 3136, 3477 (2006) (codified at 28 U.S.C. § 2241(e) (2006)).

203 *Id.* § 1405(e)(2) (codified at 10 U.S.C. § 801 (2006)).

204 548 U.S. 557, 584 (2006).

at Guantánamo Bay, including those filed before the enactment of the DTA and the MCA.[205]

In *Boumediene v. Bush*, the Supreme Court held that the MCA's elimination of habeas corpus review for Guantánamo Bay detainees violated the Suspension Clause.[206] In the wake of *Boumediene*, government lawyers faced a two-front battle: The first front was review by CSRTs followed by the D.C. Circuit, as authorized by Congress in the DTA; the second was habeas review pursuant to procedures that had yet to be defined, as opened up by the Court in *Boumediene*. Confronted by this daunting prospect of dual proceedings, and having suffered additional setbacks with respect to the scope of DTA review, government lawyers developed an inseverability-based escape argument. They argued that Congress did not intend for the judiciary to undertake DTA review—the only form of judicial review that Congress had authorized—except as a replacement for habeas corpus review eliminated in the same statute.[207] A unanimous panel of the D.C. Circuit agreed.[208] This tactic successfully terminated the DTA review process, in which the detainees had made substantial inroads against the government, and consigned the detainees to exclusive reliance on a habeas process whose contours were still a work in progress.

The D.C. Circuit's decision in *Bismullah* reveals a conventional application of existing severability doctrine.[209] The court found that "the text of the relevant provisions and the enactment of successive jurisdiction-stripping provisions demonstrate clearly that the Congress would not in the DTA have given [the D.C. Circuit Court] jurisdiction to review CSRT determinations had it known its attempt to remove the courts' jurisdiction over habeas petitions would fail."[210] According to the panel opinion, the "basic objective" of the DTA was "to restrict

---

[205] Military Commissions Act of 2006, Pub. L. No. 109-366, § 7, 120 Stat. 2600, 2635–36 (2006) (codified at 28 U.S.C. § 2241(e)), *invalidated by* Boumediene v. Bush, 128 S. Ct. 2229 (2008).

[206] 128 S. Ct. at 2274.

[207] *See* Bismullah v. Gates, 551 F.3d 1068, 1070 (D.C. Cir. 2009) ("[T]he Government argues the Congress did not intend DTA § 1005(e)(2), which gave this court alone jurisdiction to review CSRT determinations, to stand apart from the section of the Military Commissions Act (MCA) that provides no court shall have jurisdiction to hear a detainee's petition for a writ of habeas corpus . . . .").

[208] *Id.* at 1075 ("[W]e are confident the Congress would not have enacted DTA § 1005(e)(2) in the absence of the statutory provision banning the courts from exercising jurisdiction over a detainee's habeas petition. Because the latter provision has been held unconstitutional, the former must also fall.").

[209] *See id.* at 1070 ("If it is evident the Congress would not have enacted one statutory provision had it known that another provision would be held unconstitutional, then the former provision cannot be severed from the latter, and the two provisions must fall together.").

[210] *Id.* at 1072.

judicial review of the Executive's detention of persons designated enemy combatants."[211] In light of this understanding of Congress's objective, the court concluded that Congress "undoubtedly" would not have enacted the jurisdiction-granting provision of the DTA (for D.C. Circuit review of CSRT determinations) had it known that its attempt to eliminate habeas review would be held unconstitutional. The court reasoned that congressional creation of "an additional and largely duplicative process by which a detainee could challenge his detention in the court of appeals" is inconsistent with the basic objective of limiting detainee access to the courts.[212]

The D.C. Circuit's elimination of DTA review appears unobjectionable in its application of the existing approach to partial unconstitutionality. Relying on declarations filed by the Directors of the CIA and the FBI, the court noted that litigation resulting from detainees' pursuit of parallel avenues of review would "be duplicative and [would] greatly burden the Government's capacity to produce sensitive evidence."[213]

Although the D.C. Circuit's imaginary reconstruction of Congress's intent in creating DTA review did not rest on far-fetched speculation, the decision nevertheless remains troubling. The problem lies less in the ultimate result of eliminating duplicative proceedings than in the way in which the government achieved this result—by petitioning the judiciary rather than working through the politically accountable branches. While government lawyers shouldered a heavy burden in defending parallel claims, judicial lightening of this burden forestalled the need for legislative accomplishment of the same task and further centralized the formulation of detainee review policy in the judicial branch. By eliminating the only form of detainee review authorized by Congress, the court left as the only avenue for detainee review the adjudication of habeas corpus petitions pursuant to procedures to be defined by the judiciary.

The D.C. Circuit could not have accomplished this result under the proposed approach of displacement without inferred fallback law. Instead, under this approach, the court would have inquired only whether DTA review was itself constitutional and could function independently of the unconstitutional elimination of habeas. The Supreme Court's decision in *Boumediene* that DTA review was not an adequate substitute for habeas does not, of course, imply that DTA review is itself unconstitutional. Indeed, as the detainees' counsel pointed out

---

211 *Id.*

212 *Bismullah*, 551 F.3d at 1075.

213 *Id.* at 1074.

to the D.C. Circuit, the Supreme Court's opinion in *Boumediene* expressly stated that "the DTA . . . process [remains] intact."[214] As the government argued and the D.C. Circuit agreed, the Supreme Court included this statement in *Boumediene* to identify "the limited extent of its constitutional holding" and make it "as clear as possible that it was not holding the review provisions of the DTA unconstitutional."[215] Moreover, there is no question that the DTA review provision was "capable of functioning independently" of the invalid provision, which eliminated habeas review for detainees at Guantánamo Bay.[216] My proposed approach would have required that DTA review be left in place in the absence of explicit legislative direction (which Congress had not, to that point, provided).

As reflection on the foregoing example suggests, my proposed approach has both benefits and drawbacks. The next section outlines some advantages of that approach, while the third and final section considers qualifications and objections.

B. Why Displacement Without Inferred Fallback Law Is Better

The "without inferred fallback law" part of the proposed displacement approach functions as a clear-statement requirement for the legislature to express its intent for a holding of partial unconstitutionality to result in anything but simple displacement. The virtues of a clear statement rule for severability doctrine have been elaborated previously by John Copeland Nagle, Mark Movsesian, and Michael Shumsky—all of whom presupposed an excision-based framework for judicial review.[217] The clear statement aspect of my proposed displacement approach is functionally equivalent to theirs, and to that extent shares its virtues: It is easy for courts to implement, it is consistent with past legislative and judicial practice, and it provides a clear rule against which the legislature can act. On top of these beneficial features, the switch from excision to displacement provides the additional advantages of explanatory simplicity and avoidance of judicial lawmaking.

---

214 *Id.* at 1071 (quoting *Boumediene*, 128 S. Ct. at 2275) (internal quotations omitted).

215 *Id.*

216 *Id.* ("The parties do not dispute that the first and second requirements for severability are met—that is, DTA § 1005(e)(2) is constitutional and could function independently.").

217 *See* Movsesian, *supra* note 15, at 73–82 (recommending textual interpretation approach to severability of statutory provisions); Nagle, *supra* note 15, at 232–46 (arguing that general principles of statutory construction should be used to interpret severability); Shumsky, *supra* note 15, at 272–75 (arguing for applying clear statement rule to question of severability).

To begin with, the displacement without inferred fallback law approach is easy for the judiciary to implement in any specific case. Take, for example, the question of what should happen to increased hard-money contribution limits if an accompanying ban on soft-money contributions within the same statute is held unconstitutional. Unless Congress explicitly provides in statutory text that the two should stand or fall together, the displacement approach forbids a court to do anything to the hard-money contribution limits. However, if Congress does explicitly link the two provisions, then the inseverability clause linking them is enforceable just like any other law (assuming it has no constitutional defects itself). Thus, if Congress did include an inseverability clause, the unconstitutionality of the soft-money ban would result in the elimination of the increased hard-money contribution limits as well.[218]

In addition to ease of implementation, the displacement without inferred fallback law approach is generally consistent with past judicial practice. As set out in Part II, it is consistent with the original approach to partial unconstitutionality. And while displacement without inferred fallback law does break from the current excision-based conception of judicial review, its impact on case outcomes would not be significant in a quantitative sense. The only outcomes that would change are those in which courts reached a holding of inferred inseverability under the current excision-based approach. *Booker* is the only Supreme Court decision in recent years in this category.[219] The last Supreme Court holding of inferred inseverability of a federal law prior to *Booker* was the 1982 decision in *Northern Pipeline v. Marathon Pipe Line*.[220] And to find a Supreme Court

---

[218] For an analysis of this specific issue that yields the same result as my analysis here, but by different means, see Shumsky, *supra* note 15, at 228–31 (noting that Court would look to legislative intent to determine severability of McCain-Feingold soft-money ban section from rest of campaign finance reform bill).

[219] The Supreme Court applied a legislative intent test to hold an Executive Order inseverable in *Minnesota v. Mille Lacs Band of Chippewa Indians*, 526 U.S. 172, 190–94 (1999). The Court did not offer much reasoning to support its application of a legislative intent approach to determining the severability of an Executive Order. *See id.* at 191 ("Because no party before this Court challenges the applicability of [statutory severability] standards, for purposes of this case we shall assume, *arguendo*, that the severability standard for statutes also applies to Executive Orders.").

[220] 458 U.S. 50, 87 n.40 (1982). The Court stated:

> We cannot conclude that, if Congress were aware that the grant of jurisdiction could not constitutionally encompass this [state law contract claim] and similar claims, it would simply remove the jurisdiction of the bankruptcy court over these matters, leaving the jurisdictional provision and adjudicatory structure intact with respect to other types of claims, and thus subject to Art. III constitutional challenge on a claim-by-claim basis.

Id.

holding of inferred inseverability of a federal law prior to *Northern Pipeline*, it is necessary to go all the way back to the New Deal.[221] Inference-based invalidation is similarly rare in state systems.[222]

Many modern Supreme Court decisions holding federal laws partially unconstitutional are fully consistent with the displacement without inferred fallback law approach.[223] Additionally, every instance of as-applied adjudication to which the excision approach attributes implicit severance analysis is consistent with the displacement without inferred fallback law approach.

There are a number of cases that use standard severability reasoning to reach outcomes consistent with my proposed displacement approach. For this category of cases, shifting to the proposed displacement method would eliminate the case-by-case, provision-by-provision analysis that characterizes existing doctrine. If there were reason to believe that the search for legislative intent outside of explicit fallback law would actually prove fruitful in such cases, the loss of accuracy would count against the displacement approach as a cost of the change. But there is no persuasive reason to believe that this is so. To the contrary, preference-driven manipulation of severability doctrine by the judiciary is an acknowledged problem.[224]

---

221 *See* Nagle, *supra* note 15, at 220 ("The Supreme Court has not invalidated an entire federal statute as nonseverable since the 1930s."). This statement excludes *Northern Pipeline*, presumably because it treats the unit of analysis as the entire statute—in that case, the Bankruptcy Act of 1978. *See* N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 87 (1982). The holding of inseverability in *Northern Pipeline* was of application inseverability of a statutory provision, as in *Booker*. *See id.*

The last major burst of commentary and analysis surrounding severability doctrine occurred in the wake of the Court's 1983 invalidation of the legislative veto in *INS v. Chadha*, 462 U.S. 919 (1983). The circuits split on the severability of the legislative veto contained in the Reorganization Act. *Compare* EEOC v. CBS, Inc., 743 F.2d 969, 973 (2d Cir. 1984) (finding legislative veto inseverable), *with* EEOC v. Hernando Bank, Inc., 724 F.2d 1188, 1192 (5th Cir. 1984) (finding legislative veto provision severable), *and* Muller Optical Co. v. EEOC, 743 F.2d 380, 388 (6th Cir. 1984) (same). The Supreme Court held the legislative veto severable in both cases that squarely addressed the question. *See Chadha*, 462 U.S. at 931–35; Alaska Airlines v. Brock, 480 U.S. 678, 697 (1986).

222 I have not undertaken to quantify the decisions on a state-by-state basis, but state severability regimes function much like the federal system: A strong presumption of severability can be overcome by judicially perceived legislative intent. *See* Dorf, *supra* note 26, at 285 (describing state law as "remarkably uniform on questions of severability," with each state's practice tracking presumption that statutes are severable "unless the party claiming nonseverability can show that: (1) severance would leave an incoherent statute, or (2) the legislature would not have enacted the statute without the invalid portion").

223 In *City of Boerne v. Flores*, 521 U.S. 507 (1997), and *United States v. Grace*, 461 U.S. 171 (1983), for example, the Court invalidated the statutes so far as they were unconstitutional, but no further. In both cases, the Court neither mentioned severance nor suggested that its holding of partial unconstitutionality should occasion inseverability analysis.

224 *See* Stern, *supra* note 15, at 101–02 ("The only general conclusion which can be drawn from . . . [an] analysis of what the Supreme Court has both said and done . . . is that

Notwithstanding the relative infrequency of inferred inseverability, not all cases are of equal weight, and the inferred inseverability holdings tend to be heavies. Both *Booker* and *Northern Pipeline*, for instance, required systemic changes in federal adjudication. Should the potential for disruption that comes from eliminating inferred inseverability in these cases with systemic impacts count against the proposed displacement approach? To the contrary: The case for political branch involvement in restructuring is strongest in precisely such consequential cases. And although there is nothing that, in theory, forecloses legislative involvement after a holding of inseverability, the effect of such a decision, in practice, can be to take away the perceived need for such involvement.[225]

Requiring the legislature explicitly to provide for any fallback law to take effect in the event of partial unconstitutionality not only provides a clear backdrop against which to legislate, but it also operates as a preference-estimating default rule. Congress often includes severability clauses in legislation and only rarely includes inseverability clauses, a pattern that "suggests that in the vast majority of cases, Congress intends its statutes to be severable and that when it does not, it says so."[226] This pattern also holds at the state level. In fact, many states have enacted statutes that set a default rule of severability.[227]

To all these benefits that would flow from a clear statement rule requiring a deviation from severability within an excision-based framework, the proposed displacement approach adds explanatory simplicity and the avoidance of judicial lawmaking. The explanatory simplicity runs along a few dimensions. First, the displacement approach makes it unnecessary to attribute implied severance analysis to all instances of as-applied adjudication and to all holdings of partial unconstitutionality from which that explicit severance analysis is currently conspicuously absent (even though conceptually required within the current excision-based framework).

---

the Court avails itself of one formula or another in order to justify results which seem to it to be desirable for other reasons.").

[225] In this respect, *Booker* and *Northern Pipeline* present a striking contrast. Whereas the Court in *Northern Pipeline* used inseverability to goad congressional action by expanding the consequences of the constitutional holding, the Court in *Booker* used a combination of severability and inseverability holdings to render unnecessary the need for congressional action. *Cf. N. Pipeline*, 458 U.S. at 88 ("This limited stay will afford Congress an opportunity to reconstitute the bankruptcy courts or to adopt other valid means of adjudication . . . ."); *Booker*, 543 U.S. at 227 ("[T]wo provisions of the Sentencing Reform Act of 1984 (SRA) that have the effect of making the Guidelines mandatory must be invalidated in order to allow the statute to operate in a manner consistent with congressional intent.").

[226] Shumsky, *supra* note 15, at 276.

[227] *See* Dorf, *supra* note 26, at 285–86 & 295 app.

Second, displacement without inferred fallback law allows the absence of law to be treated as such. Existing doctrine rests on the mistaken presupposition that the question of fallback law for a partially unconstitutional statute is *always* within the statute's "domain"—that is, that the statute supplies law for answering the question posed of it.[228] There is often no provision in a partially unconstitutional statute directing what should be done in the case of partial invalidity. By imposing an imaginative reconstruction approach to statutory construction on *every* partially unconstitutional statute, existing doctrine begs the antecedent question of whether statutory construction is to be engaged in at all. In circumstances such as *Booker*, in which a statute runs up against a constitutional problem that could not have been anticipated at the time of enactment, it is utterly fictional to treat the statute as if it contained an answer to the question of what the fallback law should be. Such a fiction would be easier to swallow if severability doctrine were inevitable—but it is not.

Third, the proposed displacement approach makes it unnecessary to posit the existence of a "severance power" possessed by the judiciary.[229] Michael Dorf has argued that, without severability, "any judicial decision finding any law unconstitutional, on its face or as applied, would call into question the entire legal code."[230] A corollary is that "a statute that has unconstitutional applications cannot be constitutionally applied to anyone, even to those whose conduct is not constitutionally privileged, unless the court can sever the unconstitutional applications of the statute from the constitutionally permitted ones."[231] Similarly, David Gans has written that "without some severance power, a court would have to invalidate a statute as a whole if even one of the statute's clauses or provisions violated the Constitution."[232] These statements make sense within an excision-based framework for judicial review but not within the framework for judicial review set forth in *Marbury* and in the federal and state case law of the surrounding decades. Inseverability is itself an artifact of modern severability doctrine, not a necessary feature of judicial review. Because the proposed displacement approach does not require positing a "severance power," it avoids the problems of

---

[228] For criticism of this presumption, see Frank H. Easterbrook, *Statutes' Domains*, 50 U. CHI. L. REV. 533, 536 (1983) ("To find that there is 'law' on a given subject is to endow the courts with authority they lacked before. It is therefore worthwhile to demand that, before courts begin the process of 'construction,' they ascertain that the legislature has conferred the power of interpretation.").

[229] *See* Gans, *supra* note 55, at 653.

[230] Dorf, *supra* note 8, at 370.

[231] Dorf, *supra* note 26, at 238.

[232] Gans, *supra* note 55, at 653.

needing to explain and justify the courts' possession of such a power, where it comes from, and how far it reaches.

The final advantage of displacement without inferred fallback law is that it avoids the routine judicial lawmaking-by-subtraction that excision requires. Within the excision approach, every holding of partial unconstitutionality requires a judicial operation beyond judicial review itself to decide what to do with the constitutional remainder: The court must either (i) make a new law consisting of the old law minus its unconstitutional parts, or (ii) expand invalidity beyond unconstitutionality. Within the proposed displacement approach, by contrast, there is no need for a court to perform any operation on a statute to prevent metastasis from partial unconstitutionality into total invalidity. The only remaining issue after a holding of partial unconstitutionality is interpretive—what, if anything, did the legislature provide? A conception of constitutional adjudication that envisions the Court making a "new law" consisting of the old law minus its unconstitutional parts whenever it uses severability doctrine to contain the effects of a constitutional ruling assigns the judiciary an avowedly legislative function. The displacement approach, by contrast, cleanly divides responsibility between the Court and Congress: The Court is responsible for enforcing the Constitution, and Congress is responsible for legislating.

By authorizing the Court to devise a scheme approximating what Congress would have wanted, even when the idea that Congress actually wanted anything is a fiction, existing severability doctrine leads to law for which neither the Court nor Congress is entirely answerable. For example, who is responsible for the federal sentencing scheme that is now in place in the wake of the *Booker* decision? Not Congress, which legislated a different system; not the Court, which only ruled as it did because of what it concluded the enacting Congress would have wanted. And while it is true that Congress can step in at any time to change the current federal sentencing scheme, any changes it wishes to make must pass through the demanding bicameralism and presentment requirements of Article I, § 7, while the Court remains free to continue refining the federal sentencing system without such constraints.

C. Qualifications and Objections

Having set forth the displacement approach and described its advantages, various qualifications and objections warrant consideration. An initial qualification relates to the ambition of the proposal. Recovering the original approach to partial unconstitutionality shifts

the understanding of judicial review in a manner that is ambitious in one respect but modest in another. The shift from excision to displacement is ambitious inasmuch as displacement is a replacement for the reigning general framework for judicial review. After such a shift, every single instance of judicial review would be understood in terms of displacement rather than excision. Notwithstanding the magnitude of the conceptual shift, however, the impact on case outcomes would be modest, as noted above.

A second qualification relates to the ease of implementing the displacement approach as well as to its theoretical underpinnings. The implementation of displacement without inferred fallback law would be relatively simple *within* any given state system or the federal system. Courts would simply stop engaging in the counterfactual speculation required by modern severability doctrine; it might even take a while before anyone noticed any change. Matters become more complicated when considering the problem of partial unconstitutionality *across* systems. Suppose inferred inseverability analysis were eliminated from federal law but not from a particular state's law. In such a situation, may a federal court reviewing a partially unconstitutional state statute use the proposed displacement approach if state law mandates the use of inferred inseverability analysis? An absolute incapacity for such analysis in federal courts would result in the potential for different outcomes depending on whether the same constitutional challenge is litigated in state or federal court.

These considerations point to a significant limitation on the arguments for the proposed displacement approach. If the arguments against inferring fallback law yielded the conclusion that such inference is categorically beyond the power of an Article III court—say, because it calls for a task necessarily legislative rather than judicial in nature—then federal courts could not engage in it at all. Yet the arguments need not be taken to establish such a sweeping proposition. After all, federal courts have been inferring fallback law for almost a century and a half. The arguments against such inference are better understood as establishing a limit on the federal courts vis-à-vis Congress with respect to the allocation of federal legislative power, rather than a categorical absence of judicial authority to infer legislative intent more generally. At its core, the argument for displacement without inferred fallback law rests on judgments about how best to implement the separation of powers in the U.S. Constitution.

To the extent that any given state seeks to mirror the federal division of separated powers, the arguments against inferred fallback law carry over to that state system. But such mirroring is not

mandatory,[233] and federal courts facing partially unconstitutional state statutes should respect state separation-of-powers schemes that empower the state judiciary to infer fallback law. Formally, the exercise in which the federal court would be engaged is no different from the state's highest-court-predictive approach required when facing any other question of state substantive law.[234]

Another qualification relates to situations in which displacement alone does not appear to provide sufficient guidance about just what is to be displaced. Consider, for instance, a state system for regulating wine sales that discriminates against out-of-state wine shippers by permitting in-state wineries to ship directly to consumers, while requiring that all other shippers must sell through in-state wholesalers, resulting in higher costs.[235] Suppose that this discriminatory treatment with respect to direct shipping is held to violate the Dormant Commerce Clause because it discriminates against out-of-state wineries.[236] In the absence of explicit guidance from the legislature, should the court cure the discrimination by closing direct shipment to in-state shippers or, instead, by opening direct shipment to out-of-state shippers?

Under current law, a court must discern unexpressed legislative intent to determine the remedy that "best furthers the legislative purposes animating the underlying statutory scheme."[237] If courts cannot engage in that sort of hypothesizing about legislative intent regarding fallback law—and they cannot under the proposed displacement approach—then what should courts do in circumstances like these? The short answer is that courts should look to constitutional law rather than unexpressed legislative purpose in formulating a remedy.[238]

---

233 *See* Helen Hershkoff, *State Courts and the "Passive Virtues": Rethinking the Judicial Function*, 114 HARV. L. REV. 1833, 1837 (2001) ("[F]unctions that seem intuitively nonjudicial in the federal system are assigned to the judicial branch in some states, and judicial officers discharge them comfortably.").

234 *Cf.* Nicholas Quinn Rosenkranz, *Federal Rules of Statutory Interpretation*, 115 HARV. L. REV. 2085, 2152 (2002) ("[S]tate statutes arrive in federal court hand in hand with their own homegrown state interpretive methodology" because "legislating interpretive methodology is, within limits, an incident of the power to legislate.").

235 *See* Beskind v. Easley, 325 F.3d 506 (4th Cir. 2003) (holding that North Carolina's discriminatory treatment of out-of-state wineries violates Commerce Clause).

236 *See* Granholm v. Heald, 544 U.S. 460 (2005) (holding that New York and Michigan's discriminatory treatment of out-of-state wineries violates Commerce Clause).

237 Evan H. Caminker, *A Norm-Based Remedial Model for Under-Inclusive Statutes*, 95 YALE L.J. 1185, 1185 (1986).

238 *See id.* at 1202–03 ("[R]ather than choose the remedy most consistent with policies underlying the statutory scheme, courts should select the remedy that best promotes values embedded in the constitutional text."); Gans, *supra* note 55, at 696 ("Severability should not be treated as a legislative question to be answered by inspecting the intent of the

Even if the Dormant Commerce Clause problem can be cured by either the court's contraction or its expansion of the favored treatment, it does not follow that constitutional doctrine is or ought to be indifferent between the two "remedial starting point[s]."[239] Dormant Commerce Clause doctrine implements a national-market norm that supports an economic system in which "every farmer and every craftsman shall be encouraged to produce by the certainty . . . that no home embargoes will withhold his exports, and no foreign state will by customs duties or regulations exclude them."[240] When the discriminatory wine-shipping scheme described above is held to violate this doctrine, courts should resort to the underlying national-market constitutional norm in devising an appropriate remedy. Thus, the court should order expansion of the direct-shipping benefit as a "remedial starting point,"[241] regardless of what the court believes the legislature would have intended. If contraction of the benefit available to in-state interests would not violate the Dormant Commerce Clause, and if the legislature prefers that approach instead, it can so provide by passing future legislation. But the court, charged with remedying the constitutional violation resulting from the system that favors in-state interests, should not try to divine the enacting legislature's unenacted intentions. Instead, it ought to advance the underlying constitutional national-market norm with its remedy.

In addition to the foregoing qualifications, displacement without inferred fallback law is subject to a number of potential objections that must be considered. I address here two of the most prominent.

First, even if the historical analysis is persuasive in establishing that the original approach to partial unconstitutionality had no need for the hypothetical intent-based analysis required by modern severability doctrine, one could object that returning to a version of that approach would be mistaken. The argument here is that severability at some point became necessary as legislation became more complex. One response to this critique is that even if legislation became more complex throughout the nineteenth century, it is not as if legislation was ever that simple. Another response is to question whether legislative intent–based excision analysis is the most appropriate response to the increased statutory complexity that has occurred. The potential for holdings of unconstitutionality to upset legislative bargains has been present from the beginning of judicial review. If protecting implicit

enacting legislature. Instead, it is a remedial question for the courts to be considered in light of structural constitutional principles, including separation of powers.").

239 Caminker, *supra* note 237, at 1201.

240 H.P. Hood & Sons, Inc. v. Du Mond, 336 U.S. 525, 539 (1949).

241 Caminker, *supra* note 237, at 1201.

legislative bargains about partially unconstitutional statutes is the province of the judicial department, why did it take five decades from *Marbury* for a state court to develop a doctrine to serve that function[242] and another two decades for the United States Supreme Court to adopt it?[243]

A second objection to displacement without inferred fallback law flips the historical perspective: If courts have been engaged in legislative intent–based excision analysis for a century and a half, can it any longer be contended that such analysis is outside the judicial function? This objection need not rest on a form of separation-of-powers adverse possession. Professor Paul Bator was correct that the judicial power is "a purposive institutional concept, whose content is a product of history and custom distilled in the light of experience and expediency."[244] With severability doctrine, however, experience has revealed that expediency has too often overtaken history and custom. The problem is not simply that the existing excision approach is unmoored from traditional understandings of the federal judicial function and the Constitution's allocation of the legislative and judicial powers, but that the modern approach is subject to manipulation. Even those who accept the inevitability of severability should be troubled by the manner in which the Supreme Court has wielded this manipulable power to invalidate civil rights laws,[245] the federal income tax,[246] and New Deal legislation,[247] among other statutes. Should we make room in the "purposive institutional concept" of the judicial power for this doctrine and the judicial role it both presupposes and perpetuates? Not when there is an alternative whose foundation in traditional understandings of the judicial function is as solid as that of the proposed displacement approach.

---

[242] *See supra* notes 144–50 and accompanying text.

[243] *See* United States v. Reese, 92 U.S. 214, 221 (1875) ("We are not able to reject a part which is unconstitutional, and retain the remainder, because it is not possible to separate that which is unconstitutional, if there be any such, from that which is not."); *see also* Packet Co. v. Keokuk, 95 U.S. 80, 89 (1877) ("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.").

[244] Paul M. Bator, *The Constitution as Architecture: Legislative and Administrative Courts Under Article III*, 65 IND. L.J. 233, 265 (1990).

[245] *E.g.*, United States v. Harris, 106 U.S. 629, 644 (1883) (invalidating statute providing remedy for abridgement of equal protection by private conspirators); United States v. Reese, 92 U.S. 214, 221–22 (1875) (invalidating voter intimidation legislation).

[246] *E.g.*, Pollock v. Farmers' Loan & Trust Co., 158 U.S. 601, 637 (1895) (invalidating federal income tax scheme).

[247] *E.g.*, Carter v. Carter Coal Co., 298 U.S. 238, 315–16 (1936) (invalidating statute providing for regulatory oversight of coal industry); R.R. Ret. Bd. v. Alton R.R., 295 U.S. 330, 360 (1935) (invalidating statute setting up compulsory retirement and pension system for interstate railroad carriers).

Conclusion

This Article's historical analysis of the first several decades of post-*Marbury* judicial review in the United States clearly demonstrates that modern severability doctrine is not necessary for a workable system of judicial review. There is therefore no need for the courts to continue the hypothetical mind-reading that modern severability doctrine requires. Courts should instead deploy a reconstructed version of the displacement approach that operated before the advent of modern severability doctrine. Depriving courts of the power to reshape partially unconstitutional statutes to fit their perception of what the legislature would have wanted better implements the Constitution's separation of powers.

In the end, metaphors matter, and excision is pernicious. As a description of what courts do to partially unconstitutional legislation, the imagery of excision underwrites a conception of the judicial role that is avowedly legislative. The acceptance of excision also effaces an understanding of how courts can deal with partial unconstitutionality in a manner properly judicial. Looking back to the original displacement approach to partial unconstitutionality points the way forward, toward restoration of review without rewriting.