UCLA LAW REVIEW

Choosing Constitutional Remedies

Eric S. Fish

ABSTRACT

When a judge finds that a statute violates the Constitution, the statute must give way. But in many cases, there is more than one way for a judge to remedy the conflict between a statute and the Constitution. And in choosing which remedy to impose, there is usually no external source of law telling the judge what to do. They alone must decide which remedy is best.

How should judges exercise this discretion? In the American tradition, it is taken for granted that judges should use restraint—they should select the remedy that disrupts the statute as little as possible. But, as this Article shows, there are two conflicting approaches to judicial restraint when choosing constitutional remedies. One approach, herein labeled "Editorial Restraint," holds that judges should assume as little power to change legislation as possible. It posits a sliding scale of judicial interventions—adding language to a statute is worse than striking down language, which is worse than striking down an application, which is worse than adopting an avoidance interpretation. The other approach, herein labeled "Purpose Preservation," focuses instead on finding the remedy that does the least damage to the legislature's goals. That remedy might involve adding language, striking down language, striking down an application, or adopting an avoidance interpretation—what matters is the intervention's substantive effect on the statute and the legislative purpose. These two approaches are manifested in relatively pure form in the laws of England (which adopts Editorial Restraint) and Canada (which adopts Purpose Preservation), while the American doctrine of constitutional remedies is an untheorized, heterodox, and often incoherent mix of both.

This Article explores several different aspects of the American doctrine of constitutional remedies, showing that it sometimes follows the logic of Editorial Restraint and sometimes the logic of Purpose Preservation (and, sometimes, that it purports to follow one but in fact follows the other). The Article then argues that Purpose Preservation is the superior approach, and ought to be explicitly embraced in the United States. This is so because Editorial Restraint relies on false distinctions—there is no meaningful difference between adding language to a statute, striking down language, striking down applications, or adopting an avoidance interpretation. These remedial categories blur together in practice, and none is a greater intrusion into the legislative sphere than any other. Purpose Preservation, while not a perfect approach, at least safeguards the principle that majoritarian legislatures' goals should determine the content of the statutes they enact.

63 UCLA L. Rev. 322 (2016)

AUTHOR

Ph.D. candidate in law, Yale Law School. For helpful comments and feedback on prior drafts, the author would like to thank his Ph.D. classmates, Guido Calabresi, Heather Gerken, Daniel Hemel, Mike Knobler, Robert Leckey, Lisa Manheim, Jon Michaels, Lisa Ouellette, Bradley Smith, and Kate Stith.

TABLE OF CONTENTS

INTRODUCTION

In 1984, New York's rape statute provided that "A male is guilty of rape in the first degree when he engages in sexual intercourse with a female . . . [b]y forcible compulsion."[1] In a separate provision, it defined "female" as "any female person who is not married to the actor."[2] Thus only a male could be a perpetrator of rape, and only a female who was not married to the rapist could be a victim. In *People v. Liberta*, a rape defendant who was not married to his victim brought a constitutional challenge to this statute in New York's highest court.[3] He argued that the rape law's gender-based discrimination and exclusion of marital rape violated the Equal Protection Clause. The New York Court of Appeals agreed on both counts. But this left the court with a conundrum: what was the appropriate remedy? Should it strike down the statute, thereby legalizing rape in New York and entitling all people previously convicted of rape to have their convictions vacated?[4] Or should it expand the statute by effectively rewriting it, making it gender-neutral and including marital rape?[5]

The conventional answer in American legal thought is that the court should exercise restraint.[6] Courts lack the democratic legitimacy of elected bodies, so they must be careful about changing the content of the legislative code. But what, exactly, does judicial restraint require in a case like *Liberta*? There are two possible accounts. One the one hand, extending the law to include female perpetrators, male victims, and victims married to the perpetrator requires adding new language to the statute. A court choosing to extend the reach of the rape law must, in effect, change the words "male" and "female" to "person" and delete the phrase "who is not married to the actor." To create new law in this fashion is a major assumption of legislative power by the judiciary, and intuitively seems like a greater encroachment on the separation of powers than striking down the law. But, on the other hand, invalidating the law creates a much larger change in the law's substantive effect. If the New York legislature's main purpose in enacting the statute was to criminalize rape (as we can safely assume that it was), striking down the statute altogether will disrupt the legislature's goals far

1. N.Y. PENAL LAW § 130.35 (McKinney 1984).
2. *Id.* § 130.00(4).
3. 474 N.E.2d 567, 569 (N.Y. 1984).
4. *See id.* at 580.
5. This is the remedy that the court ultimately chose. *See id.* at 578–80.
6. *See generally* Barry Friedman, *The Birth of an Academic Obsession: The History of the Countermajoritarian Difficulty, Part Five*, 112 YALE L.J. 153 (2002).

more than extending the statute to cover new crimes. Which remedy better limits the countermajoritarian problem? Adding to the statute and causing a limited change? Or striking down the statute and causing a major change?

This case starkly illustrates an ambiguity in the concept of judicial restraint. Prior academic and judicial commentary has assumed that restraint is univariate. The less obtrusive the remedy, the less it will disrupt the democratic process. But this assumption conflates two different issues: (i) the nature of the action the court takes to fix the statute, whether it be adding language, striking down language, striking down applications, or interpreting language, and (ii) the extent to which the court's action alters the substantive effect of the statute. Depending on which of these variables one emphasizes, there are two very different approaches one could take to selecting constitutional remedies.

First, one could focus on the remedial action that the court takes. This approach will hereinafter be labeled "Editorial Restraint." The key tenet of Editorial Restraint is that the judiciary should take on as little power to edit legislation as possible. Editorial Restraint thus posits a sliding scale of judicial interventions in the legislative process. The ideal fix for a constitutional violation is an avoidance reading, which involves the quintessentially judicial act of interpreting a statute rather than changing its language.[7] Striking down applications and striking down statutory language require the judiciary to take more editorial control over legislation, and actually adding language to a statute is the most legislative remedy of all. What matters to the advocate of Editorial Restraint is not what actually happens to the law, but how much editorial power the judge assumes when imposing the remedy. If judges can rewrite laws, the reasoning goes, they effectively become legislators.

Second, one could instead focus on how extensively the judicial remedy changes the substantive effect of the law. This approach will hereinafter be labeled "Purpose Preservation." A judge adopting this approach examines how each remedy changes the actual function of the statute, and then imposes the remedy that best furthers the statute's intended purposes. In pursuing this inquiry, the judge

7. This Article frequently discusses constitutional avoidance as an alternative to remedies like striking a provision down. Constitutional avoidance is conventionally understood as a merits issue in American law, and not a part of the law of remedies. *See* Clark v. Martinez, 543 U.S. 371, 381 (2005) ("[The avoidance canon] is a tool for choosing between competing plausible interpretations of a statutory text . . . ."). But since avoidance does the same work as constitutional remedies, namely solving a statutory violation of the Constitution, it is included in this Article's theory of constitutional remedy selection. For ease of writing and reading, the avoidance canon is referred to as a remedy throughout. For an argument that the avoidance canon should in fact sometimes be treated as a remedy, see Eric S. Fish, *Constitutional Avoidance as Interpretation and as Remedy*, 114 MICH. L. REV (forthcoming 2016).

must make decisions about which possible version of the statute is most compatible with the enacting legislature's goals. The judge should not care whether a remedy involves reinterpreting a statute, deleting language from a statute, or adding language to a statute. The key question is how to keep the statute as close as possible to what the legislature desired while solving the constitutional problem.

This Article seeks to show that the unexplored conflict between these two approaches has confused the American doctrine of constitutional remedies, and that it has done so in a number of seemingly disparate doctrinal areas. This Article also argues that Purpose Preservation should be adopted as the governing philosophy for the selection of constitutional remedies, and that Editorial Restraint should be (for the most part) rejected. The argument proceeds in five Parts.

Part I makes a preliminary point—that when a court remedies a statutory violation of the Constitution, it necessarily alters the statute's content so as to make the statute constitutional. In so doing, Part I refutes an important scholarly misframing of the act of judicial review: the theory, advocated by John Harrison, Laurence Tribe, and Kevin Walsh, that to strike part of a statute down is merely to interpret that statute in light of the Constitution.[8] According to these scholars' view, when a judge strikes down a statutory provision as unconstitutional, the judge is not affirmatively invalidating that provision, but is only passively recognizing that the Constitution supersedes it. The judge is thus not changing the law, but simply stating what the law already is. This view is incompatible with the reality that judges in many cases have multiple options for resolving a conflict between a statute and the Constitution. If the constitutional problem could be remedied in more than one way, then the judge necessarily chooses to apply one remedy and not another.

Part II describes the two approaches to choosing constitutional remedies: Editorial Restraint and Purpose Preservation. It shows how they are each reflected in the legal systems of England and Canada, respectively. In England, courts cannot strike down laws at all—they can only interpret statutes to avoid a conflict with higher law, or if that proves impossible, make a nonbinding declaration that such a conflict exists.[9] This is a pure form of Editorial Restraint, as it walls the courts off from changing the legislative code. In Canada, by contrast, courts are instructed to choose the remedy that best fits with the legislative purpose, whether it be adding language, striking down language, striking down an application, or

8. *See* John Harrison, *Severability, Remedies, and Constitutional Adjudication*, 83 GEO. WASH. L. REV. 56 (2014); Laurence H. Tribe, *The Legislative Veto Decision: A Law by Any Other Name?*, 21 HARV. J. ON LEGIS. 1 (1984); Kevin C. Walsh, *Partial Unconstitutionality*, 85 N.Y.U. L. REV. 738 (2010).

9. *See* Human Rights Act 1998, c. 42, §§ 3–4 (UK).

adopting an avoidance reading.[10] This is a pure form of Purpose Preservation, as it ignores the nature of the remedy and focuses instead on what the legislature would have wanted. Part II also explores a third possible reason for preferring one constitutional remedy over another—Evan Caminker's argument that courts should choose the remedy that better advances the relevant constitutional norms.[11] There are multiple ways of viewing this approach: for example, as a limited rule providing that a judge should not choose a remedy that undermines the very constitutional provision being vindicated, or as a broader rule providing that a judge's choice should be guided by inchoate norms stemming from the Constitution (or from other legal sources).

Part III analyzes the American doctrine of constitutional remedies. It shows that the United States has a far less coherent approach than do either England or Canada. Here, there is no statute or judicial opinion that lays out a consistent framework. Considerations of both Editorial Restraint and Purpose Preservation inform the U.S. Supreme Court's doctrinal choices, but the Court does not distinguish these approaches or acknowledge the conflict between them. Part III explores five different types of cases that involve judges choosing between constitutional remedies, and shows how the governing law of each is motivated by this unexplored conflict between Editorial Restraint and Purpose Preservation: (1) cases where the court can remedy unconstitutional discrimination by either leveling a statute up to apply to the excluded parties or leveling it down to apply to no parties, (2) cases where a statute is held partly unconstitutional and the court must decide whether the rest of it is inseverable, (3) cases where the court can either strike down the entire statutory provision or strike down a specific application, (4) cases where a constitutional avoidance interpretation is possible, and (5) cases where the court can solicit the legislature to craft a remedy itself. Prior academic work has treated these issues as distinct and *sui generis*, but this Article reveals them all to be governed by the same internal conflict in the concept of judicial restraint.

Part IV makes a prescription: America should become more like Canada, less like England. We should embrace Purpose Preservation and reject Editorial Restraint. This is so largely because the hierarchy of remedies posited by Editorial Restraint cannot survive critical scrutiny. The difference between adding language to a statute and taking language away is simply a matter of arbitrary

10. *See* Schachter v. Canada [1992] 2 S.C.R. 679, 695 (Can.); Haig v. Canada (1992), 17 C.H.R.R. D/226 (Can. Ont. C.A.); Danielle Pinard, *A Plea for Conceptual Consistency in Constitutional Remedies*, 18 NAT'L J. CONST. L. 105, 112 (2006).

11. *See* Evan H. Caminker, Note, *A Norm-Based Remedial Model for Underinclusive Statutes*, 95 YALE L.J. 1185, 1185 (1986).

legislative drafting decisions. In the *Liberta* case, for example, the court only had to add language because the statute provided that "a male" can be guilty of rape of "a female."[12] Had the statute instead been written in gender-neutral terms, and explicit textual exceptions been made for female defendants and male victims, the court could simply have cured the constitutional problem by striking down these exceptions. Similarly, striking down applications and adopting avoidance readings cannot meaningfully be distinguished from adding or subtracting statutory language. All of these remedies involve judicial assumption of a legislative power—the power to change the meaning of a statute—and there is no sense in which one kind of remedy is any more legislative than another. In turn, the benefit of Purpose Preservation is that it minimizes interference with democratic politics by ensuring that the post-remedy statute is as close as possible to the enacting legislature's preferences. The legislature's preferred remedy is sometimes indeterminate, which does allow judges to engage in some motivated reasoning. But, all told, Purpose Preservation is still the best available approach to choosing constitutional remedies.

Finally, Part V refines the concept of Purpose Preservation and presents several caveats. First, it argues that courts should focus on the views of the enacting legislature, rather than the current legislature. This is so because the current legislature is unlikely to have ascertainable preferences with respect to the statute, and because focusing on the current legislature undermines what could be called "intertemporal democracy"—the idea that different majority coalitions get to govern at different times, and that one majority losing an election is not a repudiation of everything that majority enacted. Second, it shows that the case for Purpose Preservation is weaker in two contexts: when federal courts review state statutes, and when the legislature's preferred remedy would undermine the very constitutional right being protected. Third, it shows that Purpose Preservation is uniquely problematic when applied in the severability context, because severability decisions are not limited by the need to solve a discrete constitutional violation. Finally, it argues that courts should not deliberately choose a remedy that the current legislature will dislike so as to prompt a legislative override. Such attempts often fail, and they also undermine the principle of intertemporal democracy by taking away the power of both the enacting and the current legislature to define its own agenda.

12. *See* N.Y. PENAL LAW § 130.35 (McKinney 1984).

I. DISCRETION IN CONSTITUTIONAL REMEDIES

There are many kinds of cases where a court can choose how to fix an unconstitutional statute. Imagine, for example, that provision A and provision B are each individually valid, but are unconstitutional in combination. The reviewing judge can remedy the constitutional violation by striking down either A or B, and the choice of which provision to strike is essentially a matter of judicial discretion.[13] The exercise of this discretion is subject to review by appellate courts, which articulate the principles that govern how judges should decide such cases. But generally no external source of law—the Constitution, statutes, or anything else—dictates how the judiciary should decide between A and B. In this sense, choosing a constitutional remedy is much like imposing an equitable remedy in a private civil case.[14] The judge has discretion over which remedy to impose, and that discretion is guided only by the principles articulated in higher courts. This point seems simple, but it has profound implications for our understanding of judicial review. It means that judicial review is not a process of judges passively recognizing that statutes in conflict with the Constitution are already automatically invalid. Instead, it is a process of judges affirmatively altering statutes so that they are made consistent with the Constitution.

Three prominent legal scholars have staked out a contrary position, arguing that finding a provision unconstitutional does not involve editorial intervention by judges. In *The Legislative Veto Decision: A Law by any Other Name?*, Laurence Tribe proposes an understanding of constitutional review whereby courts do not strike down unconstitutional provisions at all. Instead, he argues that courts should be viewed as enforcing the entire statute in question (unconstitutional part included) *plus the Constitution*.[15] Under Tribe's theory, a court engaged in judicial review does not actually change a statute to eliminate the problematic provision—the court merely recognizes that the problematic provision cannot be enforced because the Constitution already trumps it.[16] In *Partial Unconstitutionality*, Kevin Walsh expands on this theory.[17] Walsh argues that judicial review

13. "Discretion" in this context does not mean that judges can choose remedies arbitrarily, but rather that they must make principled and consistent decisions as to remedy, in situations where the relevant statute and other legal materials are indeterminate, based on guiding principles that are ultimately selected by the judges themselves. *See* H.L.A. Hart, *Discretion*, 127 HARV. L. REV. 652 (2013).
14. *See* 27A AM. JUR. 2D EQUITY § 73 (2008).
15. Tribe, *supra* note 8, at 25.
16. *Id.* at 26; *see also* John Copeland Nagle, *Severability*, 72 N.C. L. REV. 203, 228–29 (1993) (explaining Tribe's position).
17. *See* Walsh, *supra* note 8.

should not be conceived as judicial "excision" of the unconstitutional provision, but as automatic "displacement" of that provision by the Constitution itself.[18] This "displacement" framing solves the problem of severability, Walsh contends, because it denies judges editorial control over statutes, casting them as mere surveyors of the legal landscape. As Walsh puts it, judicial review is "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."[19] Constitutional review thus consists of judges looking at the Constitution, looking at the statute, finding an inconsistency, and noting that the inconsistent parts of the statute are invalid (indeed, were never valid in the first place). The basic intuition behind this "displacement" framing is that even though Congress thought it enacted a statute containing A and B, if B violates the Constitution then Congress actually enacted A without B. The judicial role is merely to figure out that B was displaced. Judges do not actually delete anything—the Constitution does the deleting, judges just observe and report when that has happened. And in *Severability, Remedies, and Constitutional Adjudication*, John Harrison similarly argues that "invalidation" by courts is merely "a figure of speech," and that courts do not actually invalidate unconstitutional statutes.[20] Rather, judicial review involves courts recognizing the "Constitution's hierarchical superiority to other law," and taking this into account while deciding particular legal claims.[21] Just as in Walsh's displacement theory, judges determine where the Constitution supersedes a statute, but they do not actually do anything to the statute itself. Consequently, judicial review is simply a matter of interpretation, and not a matter of remedy.

To see why this position is untenable, let us return to the hypothetical choice between striking down provisions A or B. The implication of the displacement theory is that if the reviewing court chooses to strike down A, then it is merely interpreting the existing legal materials (the statute plus the Constitution) such that A is not good law. The statute was passed subject to the Constitution, and the reviewing court is observing that A has been automatically displaced by the Constitution. If the court instead had chosen to strike down B, however, that would mean that B was automatically displaced instead of A. But this is impossible. If the Constitution automatically displaces inconsistent statutory language, and a judge's job is merely to figure out when this has happened, then there is no room for judicial discretion in choosing which language gets displaced. It is incoherent to posit that whatever choice the judge makes has already been imposed by the

18. *Id.* at 742.
19. *Id.* at 779.
20. Harrison, *supra* note 8, at 82.
21. *Id.*

Constitution from the moment the statute was enacted. Judges are not the Terminator—they cannot make things disappear from the past. Discretion necessarily exists where there is more than one way to fix an unconstitutional statute.[22] And if judges exercise discretion in choosing whether the Constitution has displaced A or B, then there is no meaningful distinction between displacement and excision.

A possible variant of the displacement theory would hold that judges interpret the statute itself when choosing a remedy. That is, judges view the relevant statute as containing implicit instructions concerning which remedy should be chosen, and so selecting a remedy is a matter of figuring out those instructions. Thus judges' discretion is removed by positing that the statute itself has already selected the remedy (rather than the Constitution, per the displacement theory). A legislature could, of course, specify ex ante which remedy it wants imposed if the statute is found unconstitutional. For instance it could enact language stating, "if this law is found unconstitutional, provision B shall be invalid and provision A shall remain in force."[23] But in the overwhelming majority of cases, the statute is silent on the specific remedial choice that a judge faces. In such cases, judges would have to read very specific instructions into this legislative silence. It is a bit extravagant to call the divination of such instructions "interpretation" when it proceeds from a vacuum. A court could, of course, take its best guess as to which remedy the legislature would hypothetically have wanted had it addressed the issue. But the legislature does not enact its hypothetical intentions *sub silentio*. Such an exercise would have judges imaginatively reconstruct the law by putting themselves in the place of the legislature, not interpret the actual statute the legislature produced.[24] In cases where the legislature declines to tell judges which remedy to choose, it effectively delegates a legislative power to the judiciary, letting judges decide how the statute is to be fixed.[25]

22. *Cf.* Ruth Bader Ginsburg, Address, *Some Thoughts on Judicial Authority to Repair Unconstitutional Legislation*, 28 CLEV. ST. L. REV. 301, 317 (1979) ("When the court passes on the constitutionality of a statute in cases like *Westcott*, it concludes its essentially judicial business. If it declares the statute unconstitutional as written, the remaining task is essentially legislative.").

23. *See, e.g.*, Heckler v. Mathews, 465 U.S. 728, 734 (1984); Michael C. Dorf, *Fallback Law*, 107 COLUM. L. REV. 303, 304 (2007); Eric S. Fish, *Severability as Conditionality*, 64 EMORY L.J. 1293, 1332–33 n.138 (2015). For an example of such a provision, see 42 U.S.C. § 402(n) (1982). Remedial instructions might also be provided through more general language. *See, e.g.*, Organized Crime Control Act of 1970, Pub. L. No. 91-452, 84 Stat. 922, 947 (1970) (providing that RICO should be "liberally construed to effectuate its remedial purposes").

24. *See* Fish, *supra* note 23, at 1298.

25. *Cf.* Dan M. Kahan, *Lenity and Federal Common Law Crimes*, 1994 SUP. CT. REV. 345, 347 (1994) (arguing that ambiguities in criminal statutes are delegations of criminal lawmaking power to the judiciary); Thomas W. Merrill & Kristin E. Hickman, Chevron*'s Domain*, 89 GEO. L.J. 833, 833

This distinction between interpreting the law and changing the law is not merely academic. It has practical consequences in real cases. Consider again *People v. Liberta*, in which New York's highest court expanded the state's rape statute to include marital rape.[26] What did this ruling mean for defendants who had previously committed marital rape? Were they now criminally liable, or were they protected by the constitutional prohibition on ex post facto criminal punishment?[27] If the displacement theory is correct, then the decision to expand the law poses no ex post facto problem—the rape statute was enacted subject to the Equal Protection Clause and the Constitution automatically leveled the statute up at the moment of enactment. However if the court in *Liberta* actually changed the existing law rather than merely interpreting it, then the prohibition on marital rape can only be applied prospectively.[28]

Next, consider *United States v. Booker*.[29] In that case, the U.S. Supreme Court struck down several provisions of the Sentencing Reform Act in order to make the Federal Sentencing Guidelines nonbinding. However the Court did not make this remedy retroactive, meaning that defendants sentenced under the prior, binding sentencing guidelines regime could not bring challenges to their sentences.[30] If the displacement theory were correct, the Court would not have had the option to make its ruling non-retroactive. If the Court is merely observing that the Constitution has automatically displaced portions of the Sentencing Reform Act, it cannot then choose when this displacement happened.[31] Judges' role in the displacement theory is to declare what the law already is, not to decide how and when the law will change. Finally, consider the Court's decision in *Northern Pipeline Construction Co. v. Marathon Pipe Line Co*.[32] In that case the Court struck down the Bankruptcy Reform Act, declaring that bankruptcy courts as then constituted violated Article III. The Court not only made this ruling

(2001) (arguing that *Chevron* deference preserves the delegation of lawmaking authority to agencies in situations where a statute is ambiguous).

26. 474 N.E.2d 567 (N.Y. 1984).
27. *See* Marks v. United States, 430 U.S. 188, 192 (1977).
28. A constitutional remedy might also qualify as a Fifth Amendment taking under this logic if the remedy deprived someone of property. *See* Stop the Beach Renourishment, Inc. v. Fla. Dep't of Envtl. Prot., 560 U.S. 702, 715 (2010).
29. 543 U.S. 220 (2005).
30. *See* Duncan v. United States, 552 F.3d 442, 443 (6th Cir. 2009); Guzman v. United States, 404 F.3d 139, 140 (2d Cir. 2005); Lloyd v. United States, 407 F.3d 608, 610 (3d Cir. 2005).
31. *See* Chicot Cty. Drainage Dist. v. Baxter State Bank, 308 U.S. 371, 374–75 (1940) (finding that holdings of unconstitutionality can be made non-retroactive); *see also* Lemon v. Kurtzman, 411 U.S. 192, 198–99 (1973). *But see* Harper v. Va. Dep't of Taxation, 509 U.S. 86, 98–99 (1993) (holding that new interpretations of federal law announced by a court must apply retroactively to all cases still open on direct review).
32. 458 U.S. 50 (1982).

non-retroactive, it also stayed the remedy for a number of months to give Congress an opportunity to enact a new version of the law that would fix its constitutional problems.[33] The displacement theory would not permit the Court to delay its remedy in this fashion—if the Constitution automatically displaces the unconstitutional statute, then judges do not have the power to delay that displacement. Judges' discretion over the selection of constitutional remedies thus has practical implications for the exercise of judicial review. Since judges exercise a quasi-legislative power when they edit unconstitutional statutes, their remedial choices can be delayed or made non-retroactive, and are subject to the same constitutional restrictions as other legislative changes.

II. APPROACHES TO REMEDY SELECTION

As Part I illustrated, a judge who finds a statute unconstitutional exercises discretion over which remedy is applied. But that does not mean that a judge can use this discretion arbitrarily, or to further their personal views about public policy. Judges must exercise such discretion *as judges*, that is, by treating similar cases alike in a consistent, principled manner. In the American system, this means that they follow the methods of common law adjudication. A judge provides reasons for their choice of remedy, and these reasons are tested on appeal and then adopted or rejected by higher courts. Those higher courts in turn elaborate a set of precedents that will govern the selection of constitutional remedies into the future. This Part considers two alternative approaches that higher courts might adopt. The first, "Editorial Restraint," is embodied in the English model of constitutional review. It instructs judges to select the remedy that involves assuming the least editorial control over the statute, so as to preserve as best as possible the distinction between the legislative function (writing laws) and the judicial function (interpreting and applying laws). The second, "Purpose Preservation," is embodied in the Canadian model of constitutional review. It instructs judges to select the remedy that best fits with the purposes that the legislature had when it drafted the statute in question. The conflict between these two approaches is analogous to the conflict in moral philosophy between deontology and utilitarianism. Editorial Restraint (like deontology) looks to the nature of the remedial

33. *Id.* at 87–88; *see also* Buckley v. Valeo, 424 U.S. 1, 144 (1976) (staying judgment for a period not to exceed 30 days). Though it should be noted that the possibility that the European Court of Human Rights will find a violation of the Convention gives domestic declarations of invalidity added political force. *See* Aileen Kavanagh, *What's So Weak About 'Weak-Form Review'? The Case of the UK Human Rights Act 1998*, 13 INT'L J. CONST. L. (forthcoming 2015), at 23.

action itself, and imposes certain rules that restrain judges' choice of remedy, irrespective of the consequences. Purpose Preservation (like utilitarianism) looks to the consequences of the remedial action, and permits judges to choose any remedy that best preserves the outcome the legislature sought. This Part also discusses the implications of a third variable: whether the chosen remedy conflicts with constitutional norms.

A. The English Approach: Editorial Restraint

In 1998, the British Parliament introduced judicial review of legislation into the English system through a law called the Human Rights Act. The Human Rights Act empowers British judges to enforce the European Convention on Human Rights, albeit in a limited fashion. Section 3 of the Act provides: "So far as it is possible to do so . . . legislation must be read and given effect in a way which is compatible with the Convention rights."[34] Judges are thus instructed to—wherever possible—interpret statutes so that they will not conflict with the Convention. As the parliamentary white paper explaining the Human Rights Act makes clear, this requires judges to do more than just resolve legislative ambiguities in favor of Convention rights.[35] They must strain to find an interpretation of the relevant statute that makes it compatible with the Convention, even if that interpretation conflicts with the statute's text. This mandate has led to some creative interpretations. For instance, in *Ghaidan v. Ghodin-Mendoza*, the House of Lords invoked Section 3 to interpret a statute providing tenant survivorship rights to "a person who was living with the original tenant 'as his or her wife or husband'" as also providing such rights to unmarried homosexual partners.[36] In explaining this result, Lord Nicholls of Birkenhead emphasized the expansiveness of the mandate under Section 3. He noted:

> It is now generally accepted that the application of [S]ection 3 does not depend upon the presence of ambiguity in the legislation being interpreted. Even if, construed according to the ordinary principles of interpretation, the meaning of the legislation admits of no doubt,

34. Human Rights Act 1998, c. 42, § 3(1) (UK).

35. HOME DEPARTMENT, RIGHTS BROUGHT HOME: THE HUMAN RIGHTS BILL, 1997, Cm. 3782, ¶ 2.7 (UK); *see also* JoAnne Sweeny, *Creating a More Dangerous Branch: How the United Kingdom's Human Rights Act Has Empowered the Judiciary and Changed the Way the British Government Creates Law*, 21 MICH. ST. INT'L L. REV. 301, 322 (2013).

36. [2004] 2 AC 557 (HL) 560. Note that same-sex marriage did not exist in England at the time. *See also* R v. Sec'y of State for the Home Dep't [2001] EWCA (Civ) 1698 (Eng.) (interpreting section 2 of the 1997 Crime (Sentences) Act to add a requirement that multiple offenders must constitute a public safety risk in order to be sentenced to life in prison).

> [S]ection 3 may none the less require the legislation to be given a different meaning.[37]

Since *Ghaidan*, this expansive approach to Section 3 interpretation has become a settled feature of British jurisprudence.[38] A judge is thus required to basically rewrite the problematic statute under the guise of interpreting it, so long as they do not adopt an interpretation that is "inconsistent with a fundamental feature" of the legislation.[39]

If a judge cannot find an interpretation that elides the conflict between statute and Convention, then the judge may (but is not required to) make a "declaration of incompatibility" under Section 4 of the Human Rights Act, stating that the law in question violates the Convention.[40] Such a declaration does not affect the actual validity of the law, and is not even binding on the parties to the case the judge is deciding.[41] It is merely a judicial statement that the law is in conflict with the Convention. This seems rather toothless from an American perspective—we may shudder to think what would happen here if the Supreme Court depended on legislatures to implement its constitutional holdings. But the Parliament that enacted the Human Rights Act decided to rely on politics, rather than judicial invalidation, to enforce the rights enshrined in the Convention.[42] It predicted that a declaration of incompatibility "will almost certainly prompt the Government and Parliament to change the law."[43] And this prediction has turned out to be fairly accurate. Between the enactment of the Human Rights Act in 1998 and August 2011, nineteen declarations of incompatibility were issued. Of these, fourteen were subsequently acted on through amending legislation, and four had already been addressed through amendments by the time of the declaration.[44] Only one case during that period resulted in a declaration of incompatibility that the government did not act on: *Smith v. Scott*, in which the Scottish Registration Appeal Court held that disenfranchising prisoners violated the Convention.[45]

37. Ghaidan v. Godin-Mendoza, [2004] 2 AC 557, ¶ 29.
38. *See* Kavanagh, *supra* note 33.
39. *See Ghaidan*, 2 AC 557, ¶¶ 32–33.
40. Human Rights Act 1998 c. 42, § 4 (UK).
41. *Id.* § 4(6).
42. Section 10 provides for a special fast-tracked amendment procedure to fix legislation deemed incompatible with the Convention, but that process still must go through the elected government. *See id.* § 10.
43. HOME DEPARTMENT, *supra* note 35, ¶ 2.10.
44. RESPONDING TO HUMAN RIGHTS JUDGMENTS: REPORT TO THE JOINT COMMITTEE ON HUMAN RIGHTS ON THE GOVERNMENT'S RESPONSE TO HUMAN RIGHTS JUDGMENTS, 2010–11, Cm. 8162, at 29–46 (UK).
45. Smith v. Scott [2007] CSIH 9 [56] (Scot.); *see* Ruvi Ziegler, *The Missing Right to Vote: The UK Supreme Court's Judgment in* Chester *and* McGeoch, UK CONST. L. ASS'N (Oct. 21, 2013), http://wp.me/plcVqo-zw [perma.cc/EN6W-GF7S].

Some of these amendments have concerned significant and divisive issues—for instance in 2004 Section 4 was employed against a law allowing preventive detention of suspected terrorists, and Parliament responded by amending the law.[46] Declarations of incompatibility have thus proven relatively effective despite lacking any binding force, because they put political pressure on Parliament to amend legislation.

England's system of judicial remedies under the Human Rights Act is strictly sequenced. Under Section 3, a court must try to find an interpretation of the relevant statute that is consistent with the Convention. If this proves impossible, then under Section 4 the court may declare the statute incompatible with the Convention. But courts in England lack any authority to invalidate legislation. In explaining the decision not to permit judicial invalidation, Parliament emphasized the importance of a democratic mandate.[47] It thus relied on a set of intuitions about what kinds of actions are appropriate for Parliament, and what kinds of actions are appropriate for courts. Parliament makes public policy decisions that are embodied in legislation, and its mandate to write that legislation stems from democratic accountability through elections. Courts, by contrast, lack democratic legitimacy, and so are not empowered to change legislation directly. They must restrict themselves to interpreting legislation, and when they discern that a law conflicts with the Convention they can only point this out to Parliament. Thus there is a categorical preference for creative interpretation, which is considered a relatively noninvasive remedy since it involves no change to the law's text, over declarations of incompatibility, which involve the courts advocating amendments to the law.

This article refers to this approach to judicial review as Editorial Restraint, because it posits a hierarchy of judicial interventions in a statutory text, and instructs judges to make the intervention that requires assuming the least editorial power over the statute. Avoidance interpretations do not require changing the statute's text at all—they are merely judicial statements of what the text means.[48] A declaration of incompatibility requires the judge to come to a view on what the text of the legislation should be, and to express that view to Parliament (which in turn, is very likely to revise the text). A declaration of incompatibility thus, in the logic of the English model, involves taking more editorial control than does an

46. A & Others v. Sec'y of State for the Home Dep't [2004] UKHL 56 [7]; *see* RESPONDING TO HUMAN RIGHTS JUDGMENTS, *supra* note 44, at 39.

47. HOME DEPARTMENT, *supra* note 35, ¶ 2.13.

48. An avoidance interpretation might change the actual substance of the law a great deal, as discussed *infra* Part III.D. But since it does not involve adding or invalidating language, it leaves the statutory text formally unchanged.

avoidance interpretation, hence the mandatory preference for Section 3 over Section 4.[49] Worse still would be permitting a judge to strike down legislation. That would entail the judge changing the actual terms of the statute by declaring certain provisions invalid. England does not allow such judicial invalidation, as it would be undemocratic and infringe on the powers of Parliament.[50] The British have thus built a practice of judicial review that strictly distinguishes between different kinds of remedial actions, and enforces a categorical preference for the action that gives judges the least power to change the statute. Judges must interpret the problem away if they can, can only point out a conflict if no interpretation will avoid it, and can never invalidate a statute of their own authority. This approach focuses on the nature of the court's actions and posits general rules for when each kind of remedy is available, irrespective of the consequences in any particular case. The content of those rules is determined by intuitions about the proper roles of courts and legislatures, namely that the former interpret and enforce laws, while only the latter may edit them.

England's system models Editorial Restraint quite well. But the concept can also be extended beyond England's system, with its more limited role for the judiciary in rights enforcement, and into systems that permit stronger forms of judicial review. To return to *People v. Liberta*, it would be less problematic from an Editorial Restraint perspective if the New York Court of Appeals struck down New York's rape statute than if it rewrote the statute to make it gender neutral and to include marital rape.[51] This is because writing new language into a statute requires assuming more legislative power than merely invalidating the statute.[52] Striking down language is an exclusively negative action, and can only undo what the legislature has done, while being able to add new language gives judges total creative control over legislation. Editorial Restraint can also help make sense of

49. It is likely that declarations of incompatibility only seem like such a large assertion of editorial power in England because Parliament almost always acts on them. If we imagined that Parliament routinely ignored declarations of incompatibility, it would be easier to see them as involving less judicial assumption of legislative power than do Section 3 interpretations. It should also be noted that English judges have a long tradition of interpreting statutes in ways that conflict with the plain text. *See* John F. Manning, *Textualism and the Equity of the Statute*, 101 COLUM. L. REV. 1, 29–56 (2001).

50. *See* HOME DEPARTMENT, *supra* note 35, ¶ 2.13.

51. *See supra* notes 1–5 and accompanying text.

52. *See* Virginia v. Am. Booksellers Ass'n, 484 U.S. 383, 397 (1988); Blount v. Rizzi, 400 U.S. 410, 419 (1971); RORY LEISHMAN, AGAINST JUDICIAL ACTIVISM: THE DECLINE OF FREEDOM AND DEMOCRACY IN CANADA 141–42 (2006); David M. Bizar, Comment, *Remedying Underinclusive Entitlement Statutes: Lessons From a Contrast of the Canadian and U.S. Doctrines*, 24 U. MIAMI INTER-AM. L. REV. 121, 150 (1992); *see also* Ginsburg, *supra* note 22, at 324 ("Yes, extension does mean 'legislat[ing] a bit,' in fact, appreciating that the court is legislating seems to me the key to proper analysis of the issue."); Manning, *supra* note 49, at 58–78.

the preference for as-applied constitutional challenges over facial ones. If a court were restricted to only saying that a statute is unconstitutional with respect to the specific facts before it, the court would have far less editorial power over the law than if it could also choose to find the statute unconstitutional with respect to all possible fact patterns.[53] And the same logic applies to severability decisions. A judge must take on a certain degree of editorial power to strike down a statutory provision as unconstitutional, but the judge must assume significantly more such power to then also strike down other, perfectly constitutional provisions as inseverable.[54] With all of these remedial questions, the focus is not on the substantive outcome in the case, but instead on the kind of action the court takes to achieve that outcome.

One way to think about Editorial Restraint is to abstract away from the context of judicial review, and to ask how much it would infringe on the power of the legislature if judges were able to impose the remedy in question even in non-constitutional cases.[55] This helps one make sense of the hierarchy that the English system posits, and of the intuition that certain forms of judicial review are more legislative than others. Starting with constitutional avoidance, it would not add substantially to judicial power if judges could interpret statutes creatively—indeed, judges already do so in a variety of contexts.[56] However if a judge could generally declare that a statute does not apply to the case before it (striking down the statute as applied), that would cause a substantial blending of the legislative

53. *See* Gonzales v. Carhart, 550 U.S. 124, 168 (2007); Nat'l Endowment for the Arts v. Finley, 524 U.S. 569, 580 (1998); United States v. Raines, 362 U.S. 17, 21 (1960); Richard H. Fallon, Jr., *As-Applied and Facial Challenges and Third-Party Standing*, 113 HARV. L. REV. 1321, 1321 (2000).

54. *See* Ayotte v. Planned Parenthood of N. New Eng., 546 U.S. 320, 328–29 (2006); Regan v. Time, Inc., 468 U.S. 641, 652 (1984); David H. Gans, *Severability as Judicial Lawmaking*, 76 GEO. WASH. L. REV. 639, 663 (2008) ("The [severability] doctrine gives courts a wide-ranging power to rewrite statutes, and this regularly enmeshes the judiciary in making policy choices that are better left to the legislature."). This discussion of severability assumes a framing where judges can only strike down single provisions as unconstitutional, and one must justify their being able to declare further provisions inseverable. But the same logic applies if one's baseline is a system where judges must strike down the entire bill if a single provision is unconstitutional, and one has to justify their being able to save certain provisions. *See* Tom Campbell, *Severability of Statutes*, 62 HASTINGS L.J. 1495, 1499 (2011).

55. This thought experiment bears some similarity to Immanuel Kant's categorical imperative. It primes one's intuitions about the impropriety of certain actions by asking what would happen if those actions were universalized beyond their specific context. *See* IMMANUEL KANT, GROUNDING FOR THE METAPHYSICS OF MORALS 30 (James W. Ellington trans., 3d ed. 1993) ("Act only according to that maxim whereby you can at the same time will that it should become a universal law.").

56. *See* Pinard, *supra* note 10, at 115 ("Since statutory interpretation is what judges do, their ordinary day-to-day work, it will usually not raise issues of legitimacy."); *see also* William N. Eskridge, Jr., *Dynamic Statutory Interpretation*, 135 U. PENN. L. REV. 1479, 1538–44 (1987) (describing a number of cases in which the Supreme Court engaged in "dynamic" statutory interpretation, construing statutes contrary to the original legislative expectations).

and judicial roles—judges would be able to nullify the work of the legislature in a way analogous to jury nullification.[57] Judges would exercise even more legislative power if they could invalidate every possible application of the relevant law (by striking it down facially), or invalidate other statutory provisions not at issue in the case (by declaring those provisions inseverable). And if judges could add language to statutes in ordinary cases, then the judiciary would effectively become a second legislature. In any legal dispute, the relevant statute would be no more than the starting point for a judicial inquiry into what the parties' rights and obligations ought to be. This thought experiment reveals the logic of Editorial Restraint, and helps explain its intuitive appeal. The concern is not with whatever outcome a remedy creates in a particular case, but with what it means that judges can impose that type of remedy at all. Certain kinds of remedies seem more legislative than others, and less judicial, because they blur the distinction between what legislatures and judges can do to alter statutes. The fewer restraints that judges have when they fix unconstitutional statutes, the more those judges will appear to act as legislators.

B. The Canadian Approach: Purpose Preservation

The Canadian approach to choosing constitutional remedies is not enshrined in a statute, but rather in a decision of the Canadian Supreme Court. In *Schachter v. Canada*, a lower court found that a law providing fifteen weeks of paid leave for adoptive parents, but not for natural parents, violated the Canadian Charter of Rights and Freedoms (Canada's constitution).[58] On appeal, the Canadian Supreme Court enumerated five possible remedies for statutory violations of the Charter.[59] These were as follows: (1) "Severance," which consists of declaring only the offending provision invalid;[60] (2) "Striking down," which consists of declaring the entire statute invalid;[61] (3) "Reading down," which consists

57. Jury nullification is commonly understood as a kind of lawmaking power, whereby the community acting through the jury determines what the law ought to be. *See, e.g.*, Matthew P. Harrington, *The Law-Finding Function of the American Jury*, 1999 WIS. L. REV. 377, 433 (1999); Gary J. Jacobsohn, *Citizen Participation in Policy Making: The Role of the Jury*, 39 J. POL. 73, 76 (1977).

58. Schachter v. Canada, [1992] 2 S.C.R. 679, 680 (Can.).

59. Different sources have different remedy counts, but for the purposes of this Article, the count is assumed to be five. *See Schachter*, [1992] 2 S.C.R. at 695; *see also* Haig v. Canada (1992), 16 C.H.R.R. D/226 at 17 (Can. Ont. C.A.) ("[T]here are five remedies available under s.52."). *But see* Bizar, *supra* note 52, at 123–30 (counting three); Pinard, *supra* note 10, at 117 (counting four).

60. *Schachter*, [1992] 2 S.C.R. at 696; KENT ROACH, CONSTITUTIONAL REMEDIES IN CANADA 14-39 to 14-46 (1994).

61. *Schachter*, [1992] 2 S.C.R. at 697, 705, 715 (distinguishing between "striking down" and "severance"); *Haig*, 16 C.H.R.R. D/226 at 15; ROACH, *supra* note 60, at 14-38 to 14-39.

of interpreting the statute in such a way that it does not conflict with the Charter;[62] (4) "Reading in," which consists of extending the statute's reach to fix unconstitutional exclusions;[63] and (5) "Suspending" the remedy, which consists of selecting a remedy but delaying implementation to give the legislature a chance to craft its own solution.[64] In contrast to the English approach, the Canadian Supreme Court did not arrange these remedies in a hierarchy and require judges to only impose one kind of remedy if a different kind is not available. The Court instead acknowledged that "[a] court has flexibility in determining what course of action to take following a violation of the *Charter*."[65] Where England gives judges a fixed menu, Canada lets judges choose à la carte.

In considering what to do in *Schachter*, the Canadian Supreme Court focused on finding the remedy that the legislature would have preferred.[66] The Court thus established the principle that when a judge has the discretion to impose one remedy or another in a constitutional review case, he or she should choose the remedy that best preserves the legislature's goals. Ultimately this led the Court in *Schachter* to decide that "suspended severance" was the best remedy (rather than "reading in"), because the law in question was specifically intended to help adoptive parents, and it would stretch the law beyond recognition to extend it to the (vastly larger) category of natural parents.[67] Lower courts have followed suit. In *Haig v. Canada* the Ontario Court of Appeal considered a challenge to an antidiscrimination statute, the Canadian Human Rights Act, alleging that the Act unconstitutionally omitted homosexuals from protection.[68] The Ontario court agreed that this exclusion was unconstitutional and, applying *Schachter*, determined that the remedy that would least interfere with the legislature's objectives was to expand the statute

62. *Schachter*, [1992] 2 S.C.R. at 719–20; *see also* PETER W. HOGG, CONSTITUTIONAL LAW OF CANADA 393–94 (3d ed. 1992); ROACH, *supra* note 60, at 14-7 to 14-38.

63. *Schachter*, [1992] 2 S.C.R. at 698; *see also* PETER W. HOGG, CONSTITUTIONAL LAW OF CANADA 929 n.51 (4th ed. 1997) ("Reading *in* . . . involves adding new words to a statute to remove a constitutional defect . . . ."); ROACH, *supra* note 60, at 14-46 to 14-63.

64. *Schachter*, [1992] 2 S.C.R. at 715; ROACH, *supra* note 60, at 14-63 to 14-79.

65. *Schachter*, [1992] 2 S.C.R. at 695.

66. The Canadian Supreme Court emphasized preserving legislative goals throughout the opinion in *Schachter*. *See id.* at 697, 700, 707–09, 718.

67. *Id.* at 722–24 ("In this case, reading in would not necessarily further the legislative objective and it would definitely interfere with budgetary decisions in that it would mandate the expenditure of a greater sum of money than Parliament is willing or able to allocate to the program in question."); *id.* at 726 ("[I]t cannot be assumed that the legislature would have enacted the benefit to include the plaintiff."). The Court also noted that while the case was pending, the Canadian Parliament had implemented a legislative solution, equalizing adoptive and natural parents' leave time at ten weeks (less than the original fifteen). *Id.* at 724.

68. Haig v. Canada, (1992), 16 C.H.R.R. D/226 (Can. Ont. C.A.); *see also* Esquega v. Canada, 2008 F.C.A. 182, [2009] 1 F.C.R. 448 (Can.) (adopting "reading down" as the remedy that least interferes with legislative objectives).

by "reading in" protections for homosexuals.[69] It would have been rather ironic to strike down a broad civil rights statute so as to preserve discrimination against a group.[70] The court thus effectively ordered that the words "sexual orientation" be added to the Act.[71] The Canadian Supreme Court followed a few years later in *Vried v. Alberta,* reading sexual orientation into a different civil rights statute.[72] As these cases illustrate, when Canadian judges must choose which kind of remedy to impose—whether it be reading in, reading down, severance, invalidation, or another—they try to select the remedy that best fits with the law's purposes.

This Article refers to this approach to judicial review as "Purpose Preservation" because it is directed at ensuring that the court's remedy interferes with the law's purpose as little as possible. It is a close relative of purposivism in the statutory interpretation context. Just as purposivists resolve statutory ambiguities by choosing the interpretation that best furthers the law's goals, devotees of Purpose Preservation decide between remedies by selecting the one that best furthers the law's goals.[73] This inquiry is directed at the substantive result of a remedial order, and not at the kind of action that the court takes. Thus while Editorial Restraint focuses on the nature of the court's action, Purpose Preservation focuses instead on its consequences. Striking down language is not inherently any better or worse than adding language, adopting an avoidance interpretation, or disallowing a specific application—all that matters is finding the remedy that best serves the legislative purpose.

Of course, determining the legislative purpose can be a difficult matter. Legislatures have many different members, each with his or her own goals, and one piece of legislation can be designed to achieve multiple (sometimes conflicting) objectives.[74] There are nonetheless some easy cases. It is hard to imagine that the New York legislature would have preferred legalizing rape to criminalizing marital rape,[75] and also to imagine that the Canadian parliament would have preferred legalizing all discrimination to prohibiting discrimination against homosexuals.[76]

69. *Haig*, 16 C.H.R.R. D/226 at 18–23.
70. *Id.* at 23 ("It is therefore inconceivable to me that Parliament would have preferred no human rights Act over one that included sexual orientation as a prohibited ground of discrimination.").
71. *Id.* at 24 (ordering "that the *Canadian Human Rights Act* . . . be interpreted, applied and administered as though it contained 'sexual orientation' as a prohibited ground of discrimination in s.3 of that Act").
72. Vriend v. Alberta [1998] 1. S.C.R. 493.
73. *See* Eskridge, *supra* note 56, at 1480.
74. *See* Kenneth A. Shepsle, *Congress Is a "They," Not an "It": Legislative Intent as Oxymoron*, 12 INT'L REV. L. & ECON. 239, 248–49 (1992); Cass R. Sunstein, *Interpreting Statutes in the Regulatory State*, 103 HARV. L. REV. 405, 427 (1989).
75. *See* People v. Liberta, 474 N.E.2d 567, 578–79 (N.Y. 1984).
76. *See Haig*, 16 C.H.R.R. D/226.

But others are not so easy. Take for example the statute at issue in *Schachter*, which provided fifteen weeks of paid leave for adoptive parents but not for natural parents. It would have been incomplete for the Supreme Court to decide that this statute's overriding purpose was to ensure children are cared for, and thus that expanding the statute was the best remedy because that would mean more parental leave. The legislature also presumably had reasons for its decision to limit the statute to adoptive parents—perhaps that it would cost less money than extending it to all parents, or that adoptive parents have peculiar difficulties that can be helped by parental leave.[77] The key is to determine which remedy best serves the full matrix of objectives that the statute was enacted to further. To do so, the court must look at the available evidence—the wording of the statute, its political and legal context, its legislative history, etc.—and try to decide which remedy the legislature would hypothetically have chosen.[78] This poses a danger. In cases where the legislative will is indeterminate, judges will have to put themselves in the place of the legislature and try to decide which remedy it would have preferred. It seems likely that, in such cases, a judge's own social and political beliefs would be difficult to separate from their prognosis of what the legislature would have wanted, especially in cases where the answer is indeterminate.[79]

C. Irony Avoidance and Norm Recognition

Editorial Restraint and Purpose Preservation are both methods for respecting the legislature in constitutional review cases—the former by limiting judges' power to edit laws, and the latter by instructing judges to find the remedy the legislature would have wanted. But respect for the legislature is not the only consideration one might focus on when deciding between constitutional remedies. A judge could also try to impose the remedy that best vindicates the relevant constitutional norms. Evan Caminker defends such an approach in the

77. *See* Schachter v. Canada, [1992] 2 S.C.R. 679, 722–23 (Can.).

78. Legislation scholars distinguish between purposivism and intentionalism in statutory interpretation, with the former approach instructing judges to look to the objectively manifested purposes of the statute and the latter pointing instead to the subjective intentions of the legislature that enacted it. *See, e.g.*, AHARON BARAK, PURPOSIVE INTERPRETATION IN LAW 260–68 (2005); Eskridge, *supra* note 56, at 1480. In the context of choosing constitutional remedies, this distinction does not matter much. If one takes into account the full matrix of purposes that a law serves, then it is hard to imagine a situation where remedy A will best serve those purposes, but the legislature would instead have chosen remedy B. And even if such a case existed, the judiciary would be unlikely to know the legislature's hypothetical intentions except by looking at its general objectives. Thus the two inquiries collapse into each other.

79. *See* Bruce K. Miller, *Constitutional Remedies for Underinclusive Statutes: A Critical Appraisal of* Heckler v. Mathews, 20 HARV. C.R.-C.L. L. REV. 79, 89–90 (1985); *see also* Fish, *supra* note 23, at 1322–25.

context of cases where a statute is unconstitutionally underinclusive and judges must either strike the law down or expand it to add the excluded group.[80] Caminker argues that, in such cases, judges should be guided by "inchoate norms" embedded in the Constitution.[81] Writing in the Canadian context, Nitya Duclos and Kent Roach propose much the same approach, urging that judges should be guided by "constitutional hints" when they select remedies.[82]

In making their argument, Caminker, Duclos, and Roach conflate two different principles through which constitutional norms might determine the selection of constitutional remedies. These are worth distinguishing here. This Article will label them "Irony Avoidance" and "Norm Recognition." Irony Avoidance is the principle that when a judge fixes a constitutional violation they should not choose a remedy that will undermine the very constitutional norm being protected. For example, if a law prohibits one group of people from picketing near a school, but allows another group, and this discrimination is found to violate the First Amendment as a content-based discrimination, the court should strike down the statute rather than leveling it up, since it would be deeply ironic to prohibit picketing in order to protect freedom of speech.[83] Similarly, if a statute guaranteeing maternity leave unconstitutionally excludes fathers, the court should expand the leave benefit to men rather than denying it to women, since repealing maternity leave would undermine sex equality.[84] And, to return to the case of *Haig v. Canada*, it would be almost comical to strike down an anti-discrimination law in order to remedy the discriminatory exclusion of homosexuals from that very law.[85] In each of these examples, leveling down is a poor remedy because it creates a conflict with the constitutional norms that animate the very right being enforced. It makes sense to avoid such irony. To the extent that judges have remedial discretion they should, all else equal, exercise it to vindicate the purposes of the trumping constitutional provision. Irony Avoidance could be seen as itself

80. *See* Caminker, *supra* note 11.
81. *Id.* at 1185.
82. Nitya Duclos & Kent Roach, *Constitutional Remedies as "Constitutional Hints": A Comment on* R. v. Schachter, 36 MCGILL L.J. 1, 24 (1991).
83. *See* Caminker, *supra* note 11, at 1194–96 (discussing *Police Department of Chicago v. Mosley*, 408 U.S. 92 (1972)).
84. *Id.* at 1196–1202. *Cf.* Schachter v. Canada, [1992] 2 S.C.R. 679, 701–02 (Can.) (discussing the irony of striking down welfare benefits for single mothers on "equality" grounds). Though it should be noted that this would depend on how the relevant norm is interpreted—a judge might view the Equal Protection Clause as only advancing a norm of formal equality, and thus not find it ironic at all to eliminate maternity leave on sex equality grounds. *See, e.g.*, Reva B. Siegel, *Equality Talk: Antisubordination and Anticlassification Values in Constitutional Struggles over* Brown, 117 HARV. L. REV. 1470 (2004) (describing two very different purposes that are ascribed to the Equal Protection Clause, which each embody a distinct vision of egalitarianism).
85. *See* Haig v. Canada, (1992), 16 C.H.R.R. D/226 (Can. Ont. C.A.).

a form of purpose preservation, where the judge chooses the remedy that best preserves the purposes of the relevant constitutional right rather than the purposes of the statute.

In contrast, Norm Recognition is much broader. It involves situations where the norm being invoked comes not from the specific constitutional right being enforced, but from elsewhere in the Constitution (or another source of law).[86] One example of this is the Canadian case *R. v. Hebb*.[87] *Hebb* involved a statute that required a court, before it imprisoned a 16- to 22-year-old defendant for failing to pay a criminal fine, to produce a report on that defendant's ability to pay. The statute was challenged on age discrimination grounds by a defendant over the age of twenty-two. The reviewing judge found this challenge persuasive. The judge then opted to expand the report requirement to all defendants rather than eliminate it, on the grounds that while constitutional due process did not necessarily require an ability-to-pay determination, the practice of judicial consideration before incarceration was "constitutionally encouraged."[88] Thus the challenge in *Hebb* involved age discrimination, but the selection of remedy was driven by constitutional due process norms.[89]

Another example is *Liberta*.[90] In that case, the New York Court of Appeals determined that the state's rape statute was unconstitutionally discriminatory because it excluded female perpetrators, male victims, and marital rape.[91] Expanding the statute to include all of these cases, however, comes into strong tension with the constitutional norm of *nulla poena sine lege*—that conduct cannot be criminalized unless there is a statute prohibiting it.[92] While leveling the statute up may not technically violate the constitutional prohibition on common law crimes, the Constitution could be read to strongly discourage judicial expansion of criminal laws. In each of these cases, judges could find that the statute is invalid under one provision of the Constitution, and then decide on the appropriate remedy by looking to the norms underlying other provisions.

86. There are some borderline situations where it is not immediately clear whether or not a constitutional norm is the same as the one being enforced. For instance, if a statute is unconstitutional because it causes racial discrimination, and one of the possible remedies would undermine the norm against gender discrimination, we must decide whether these are different constitutional norms or just different manifestations of the same norm.

87. R. v. Hebb, (1989), 89 N.S.R. 2d 137 (Can.).

88. *Id.* at 151.

89. *See* Duclos & Roach, *supra* note 82, at 26–27.

90. People v. Liberta, 474 N.E.2d 567 (N.Y. 1984).

91. *See supra* notes 1–5 and accompanying text.

92. *See* United States v. Wiltberger, 18 U.S. 76, 95 (1820) (Marshall, C.J.) ("It is the legislature, not the Court, which is to define a crime, and ordain its punishment."); United States v. Hudson, 11 U.S. 32, 34 (1812).

Norm Recognition is broader than Irony Avoidance, because there are so many possible norms one could invoke. If you look through the entire Constitution for hints to guide the choice of remedy, you will commonly find multiple hints pointing in different directions. This multiplicitousness is compounded by the basic indeterminacy of "constitutional norms," which are basically abstract values ascribed to open-ended constitutional texts.[93] The fact that many different norms can be imputed to constitutional phrases like "equal protection," "due process," or "freedom of speech" leads to disagreements over which values are actually embedded in the Constitution. For example one judge might view the Equal Protection Clause as forbidding only discriminatory classifications, while another might see it as advancing substantive social equality for disadvantaged groups.[94] Similarly, one judge might view the Due Process Clause as furthering only procedural justice values, while another might see substantive due process as a constitutional value.[95] This ideologically inflected indeterminacy, combined with the possibility of conflicts between constitutional values, will lead to cases where each judge chooses a value and there is no neutral way of deciding between them. Thus, for instance, in a case where a benefit must be leveled up or down after a finding of discrimination, one judge might decide that a constitutional value favoring limited federal power favors leveling down,[96] while another might think that that a constitutional value favoring substantive equality favors leveling up.[97] Each judge could muster interpretive evidence to support his or her chosen value. It is hard to see how one could decide in a neutral way which value should trump the other.

This difficulty is further compounded by the fact that important norms can also be found outside of the Constitution. Why, indeed, do constitutional norms necessarily trump deeply entrenched legal norms that are embedded in statutes, judicial doctrines, and other sources of law? To return to *Liberta*, a

93. *Cf.* Jack M. Balkin, *Framework Originalism and the Living Constitution*, 103 NW. U. L. REV. 549 (2009) (arguing that the open texture of several major constitutional provisions permits their meaning to change); Thomas I. Emerson, *Toward a General Theory of the First Amendment*, 72 YALE L.J. 877 (1963) (providing a number of different values served by the First Amendment).

94. *See* Siegel, *supra* note 84.

95. *Compare* City of Chi. v. Morales, 527 U.S. 41, 85 (1999) (Scalia, J., dissenting), *with* Lawrence v. Texas, 539 U.S. 558, 578 (2003) (opinion by Kennedy, J.).

96. *See* Nat'l Fed'n of Indep. Bus. v. Sebelius, 132 S. Ct. 2566, 2646 (2012) (Scalia, Kennedy, Thomas & Alito, JJ., dissenting) (referring to "the background principle of enumerated (and hence limited) federal power").

97. *See* Ralph K. Winter, Jr., *Poverty, Economic Equality, and the Equal Protection Clause*, 1972 SUP. CT. REV. 41, 48–58 (describing the development of equal protection doctrine in the Warren Court, and concluding that "[a] fair evaluation of both the views of the commentators and the existing case law would be that the law is well along the road toward substantive equal protection as a vehicle of income redistribution").

strong argument could be made that the norm favoring criminalization of rape is more deeply entrenched than the norm against common law crimes.[98] It is inconceivable that a state would legalize rape, while judicially created crimes existed for much of American history.[99] And some constitutional norms are not very deeply entrenched at all. For instance the Twenty-Seventh Amendment arguably creates a constitutional norm against congressional self-dealing, but this norm is relatively weak in our legal system.[100] And many other constitutional clauses are unimportant or anachronistic.[101] Meanwhile, statutes like the Civil Rights Act,[102] the Sherman Act,[103] and the Fair Labor Standards Act[104] are sources of the most powerful and deeply embedded norms in our legal system.[105] It would seem odd for judges to invoke norms embodied in the Twenty-Seventh Amendment, but not the Civil Rights Act. And if Norm Recognition were expanded to include statutes and other sources of law, then judges could find a norm pointing in each direction in virtually any case of remedial discretion. Norm Recognition thus becomes a wide-ranging and comprehensive approach to selecting constitutional remedies. A judge following this approach would survey the entire legal-political landscape for relevant norms, determine their relative strength in relation to the remedial choices available, and select the remedy that vindicates the norm that the judge finds most important.[106] This closely resembles what Judge Hercules does in Ronald Dworkin's interpretive theory.[107] The judge first imagines that all of the existing legal structures and prior decisions

98. Here this Article is assuming that there is no federal constitutional norm favoring the criminalization of rape, though some have argued that the Thirteenth Amendment provides such a norm. *See, e.g.*, Neal Kumar Katyal, Note, *Men Who Own Women: A Thirteenth Amendment Critique of Forced Prostitution*, 103 YALE L.J. 791 (1993); Jane Kim, *Taking Rape Seriously: Rape as Slavery*, 35 HARV. J.L. & GENDER 263 (2012).

99. Indeed, some argue that federal criminal law is still effectively a type of common law. *See, e.g.*, Kahan, *supra* note 25.

100. Congress, indeed, frequently violates the Twenty-Seventh Amendment without much fuss being raised. *See* Eric S. Fish & Daniel Hemel, *Congress' Unconstitutional Pay Freeze*, 34 NAT'L L.J. 34 (2012).

101. *See, e.g.*, U.S. CONST. art. I, § 2 ("[D]irect Taxes shall be apportioned among the several States which may be included within this Union . . . ."); U.S. CONST. art. II, § 1 ("No Person except a natural born Citizen, . . . shall be eligible to the Office of President."); U.S. CONST. art. III, § 3 ("No Person shall be convicted of Treason unless on the Testimony of two Witnesses to the same overt Act, or on Confession in open Court."); U.S. CONST. amend. III.

102. Pub. L. No. 88-352, 78 Stat. 241 (codified as amended in scattered sections of 2 U.S.C., 28 U.S.C., and 42 U.S.C.).

103. 15 U.S.C. §§ 1–7 (2012).

104. 29 U.S.C. §§ 201–05 (2012).

105. *See* William N. Eskridge, Jr. & John Ferejohn, *Super-Statutes*, 50 DUKE L.J. 1215, 1225–27 (2001).

106. *Cf.* Sunstein, *supra* note 74 (defending an analogous approach in the statutory interpretation context).

107. *See* RONALD DWORKIN, LAW'S EMPIRE 337–54 (1986).

in the system taken together create a coherent hierarchy of norms, and then chooses the remedy that best fits with their normative interpretation of that hierarchy.[108] Disagreements over the choice of remedy thus become disagreements over which norms are more deeply entrenched.

III. UNRAVELING THE AMERICAN APPROACH

Turning now to the American system, for decades academics and jurists concerned with the practice of judicial review have been wrestling with the countermajoritarian problem.[109] At a basic level, it conflicts with democratic norms to have unelected (and life-tenured) judges make decisions that change laws enacted by elected legislatures. Editorial Restraint and Purpose Preservation represent two different strategies for dealing with the countermajoritarian problem in cases where a judge can choose between several remedies. Editorial Restraint lets judges conceal their discretion by asserting as little creative control over the language of the statute as possible, and thus lets judges avoid being accused of usurping the legislature's role. Purpose Preservation, by contrast, requires judges to exercise discretion openly, so as to best approximate the result that majoritarian processes would have created. These strategies are fundamentally different—Editorial Restraint involves judicial abdication of power over legislation, while Purpose Preservation requires judicial assertion of such power. Yet the American doctrine has not distinguished these approaches, and switches incoherently between them on various remedial questions. Where England and Canada each have a comprehensible, consistent principle governing how remedies should be chosen, the American approach remains untheorized.

This Part explores the American doctrine of constitutional remedies concerning five different types of cases, and shows how the doctrine is animated by the logic of Editorial Restraint, the logic of Purpose Preservation, or both.[110]

108. *See id.* at 245.

109. *See* ALEXANDER M. BICKEL, THE LEAST DANGEROUS BRANCH: THE SUPREME COURT AT THE BAR OF POLITICS 16 (1962); Erwin Chemerinsky, *The Supreme Court, 1988 Term—Foreword: The Vanishing Constitution*, 103 HARV. L. REV. 43, 71 (1989) ("Most constitutional scholars for the past quarter-century . . . have seen the task of constitutional theory as defining a role for the Court that is consistent with majoritarian principles."); Richard H. Pildes, *Is the Supreme Court a "Majoritarian" Institution?*, 2010 SUP. CT. REV. 103, 104 (2010) ("[I]t is not wrong to characterize American legal thought as 'obsessed' with the moral problem of judicial review.").

110. Unfortunately, the discussion in this Part is limited to cases involving federal courts applying the federal Constitution. While it is important to determine how state courts choose constitutional remedies, the sheer number of state systems renders such a project beyond the scope of this Article. For a survey of the fifty states' approaches to severability, see Michael C. Dorf, *Facial Challenges to State and Federal Statutes*, 46 STAN. L. REV. 235, 295–304 (1994).

First, the Supreme Court follows the logic of Purpose Preservation when extending underinclusive statutes. Second, although the Court's doctrine in severability cases calls for Purpose Preservation, in practice the Court follows the logic of Editorial Restraint. Third, the Court's doctrine concerning facial challenges disfavors them according to the logic of Editorial Restraint, but in practice the Court frequently allows them based on the logic of Purpose Preservation, Irony Avoidance, or other case-specific considerations. Fourth, the Court states a categorical preference for constitutional avoidance interpretations where they are available, à la Editorial Restraint, though in one major case—*United States v. Booker*—it asked instead whether the enacting legislature would have preferred the avoidance interpretation, à la Purpose Preservation.[111] Finally, the Court only rarely solicits a legislative remedy from Congress or the relevant state legislature.[112]

A. Leveling Up or Leveling Down

Adding language to a statute—what Canadian courts label "reading in"—is the least preferred remedy from an Editorial Restraint perspective.[113] However, within the logic of Purpose Preservation, adding language is the best option if it vindicates the legislative goals better than any alternative remedy. Generally, American courts adopt the Editorial Restraint approach when they consider whether to rewrite or add language to a statute. They often behave as though such a remedy is beyond their power. Thus the Supreme Court stated in *Virginia v. American Booksellers Association* that "we will not rewrite a state law to conform it to constitutional requirements,"[114] and in *Blount v. Rizzi* that "it is for Congress, not this Court, to rewrite the statute."[115] There is one major exception to

111. 543 U.S. 220 (2005).

112. This Part discusses situations where Purpose Preservation and Editorial Restraint both have something to say about the choice of remedy. It does not consider situations where either approach would be indifferent between remedies. For instance if a judge were to choose which of two provisions to strike down, or which of two avoidance interpretations to adopt, the Editorial Restraint approach would not have much to say about which remedy should be chosen, and so the judge would have to adopt a different theory. *See, e.g.*, Bowsher v. Synar, 478 U.S. 714, 783–87 (1986) (Blackmun, J., dissenting) (using the Purpose Preservation approach in considering which of two provisions to strike down).

113. For some scholars and commentators, adding language to statutes is tantamount to judicial legislating. *See, e.g.*, LEISHMAN, *supra* note 52, at 141; Bizar, *supra* note 52; Ginsburg, *supra* note 22, at 324.

114. 484 U.S. 383, 397 (1988).

115. 400 U.S. 410, 419 (1971); *see also* United States v. Nat'l Treasury Emps. Union, 513 U.S. 454, 478 (1995); W. Va. Univ. Hosps., Inc. v. Casey, 499 U.S. 83, 101 (1991); Am. Tobacco Co. v.

this approach, however. In equal protection cases, the Court will often level the statute up to include the unconstitutionally excluded group.[116] It will do so even in cases where there is no textual exception to strike down, and leveling up thus requires adding language to the statute.

This special treatment for equality cases began in the 1970s.[117] Concurring alone in *Welsh v. United States*, Justice Harlan first distinguished equality cases from other kinds of constitutional review cases, arguing that it is preferable to level a statute up if striking it down would frustrate the legislature's goal to benefit third parties.[118] He justified such an approach through a Purpose Preservation rationale, arguing that the courts enjoy a "presumed grant of power . . . to decide whether it more nearly accords with Congress's wishes to eliminate its policy altogether or extend it in order to render what Congress plainly did intend, constitutional."[119] This approach took off as the Court began applying the Equal Protection Clause to laws that discriminated by sex. In *Frontiero v. Richardson*,[120] *Weinberger v. Wiesenfeld*,[121] *Califano v. Westcott*,[122] and other cases,[123] the Court considered challenges to social security provisions and other benefits laws that provided greater benefits to women than to men, and in each case it expanded the benefits for men to achieve gender parity. *Weinberger*, for example, involved a statute that provided social security payments to widows with dependent children, but not to widowers. The Court remedied this sex discrimination by writing widower fathers into the statute.[124] In an article written in 1979, then-professor Ruther Bader Ginsburg (the main architect of the Court's sex equality

Patterson, 456 U.S. 63, 75 (1982); Aptheker v. Sec'y of State, 378 U.S. 500, 515 (1964); Hill v. Wallace, 259 U.S. 44, 70–71 (1922); United States v. Reese, 92 U.S. 214, 221 (1875).

116. *See* Bizar, *supra* note 52, at 130–31.

117. Two prior cases, *Iowa-Des Moines National Bank v. Bennett*, 284 U.S. 239 (1931), and *Skinner v. Oklahoma*, 316 U.S. 535 (1942), raised the question of leveling up in equality cases. But the Court did not level up in either case, nor did it discuss the issue at much length or consider the relevance of legislative purpose. *See* Ginsburg, *supra* note 22, at 306–10.

118. 398 U.S. 333, 366 n.18 (1970) (Harlan, J., concurring in result). Justice Harlan emphasized the existence of a broad severability clause in that case, arguing that it empowered the Court to expand the statute. *Id.* at 364–65.

119. *Id.* at 355–56.

120. 411 U.S. 677 (1973).

121. 420 U.S. 636 (1975).

122. 443 U.S. 76, 93 (1979).

123. For a more comprehensive list, see Candice S. Kovacic, *Remedying Underinclusive Statutes*, 33 WAYNE L. REV. 39, 49–50 nn. 66–72 (1986). Some state court decisions extending statutes after constitutional equality challenges predated the 1970s. *See, e.g.*, Burrow v. Kapfhammer, 145 S.W.2d 1067 (Ky. 1940); Quong Ham Wah Co. v. Indus. Accident Comm'n of Cal., 192 P. 1021 (Cal. 1920).

124. *See Weinberger*, 420 U.S. at 653, *aff'g* 361 F.Supp. 981 (D.N.J. 1973); Ginsburg, *supra* note 22, at 302–03.

jurisprudence) defended leveling up in these cases. She argued that "The courts act legitimately . . . when they employ common sense and sound judgment to preserve a law by moderate extension where tearing it down would be far more destructive of the legislature's will."[125] The Court fully adopted this rationale in *Heckler v. Mathews*, articulating a full-throated endorsement of Purpose Preservation in equal protection cases.[126] This embrace of Purpose Preservation has been extended by lower courts to other constitutional equality guarantees including the Dormant Commerce Clause and the First Amendment.[127]

The American doctrine on adding language to statutes has two different logics. In most cases, courts do not permit themselves to add language. They cannot, for instance, add new procedures to a statute to satisfy due process requirements, grant bankruptcy judges life tenure to satisfy Article III, or rewrite impermissibly vague statutes to make them more specific.[128] But, for constitutional equality cases, courts do add language to a statute if doing so would better preserve the legislative purpose. So the doctrine embraces Editorial Restraint in general, but Purpose Preservation for equality cases. How to account for this distinction? Justices Harlan and Ginsburg each justify extension remedies in equal protection cases by arguing that they better vindicate the legislative will.[129] But this does not show why equality cases are unique—why are courts barred from rewriting statutes following due process or vagueness challenges, if doing so better vindicates the legislative will? It is difficult to find a principled distinction here, but a few features of equality cases seem to make them especially suitable to extension remedies. First, there is generally a clear, determinate, defined number which can be leveled up to—for example, if women receive ten weeks of parental

125. Ginsburg, *supra* note 22, at 324; *see also* Deborah Beers, *Extension Versus Invalidation of Underinclusive Statutes: A Remedial Alternative*, 12 COLUM. J.L. & SOC. PROBS. 115 (1975) (showing that the Court focuses mainly on legislative intent when deciding whether to extend or invalidate).

126. 465 U.S. 728, 739 n.5 (1984).

127. *See, e.g.*, Freeman v. Corzine, 629 F.3d 146, 164 (3d Cir. 2010) (noting the extension of Purpose Prevention to the dormant Commerce Clause); Wiesmueller v. Kosobucki, 571 F.3d 699, 702 (7th Cir. 2009) (same); Rappa v. New Castle Cty., 18 F.3d 1043, 1073 (3d Cir. 1994) (noting the extension of Purpose Prevention to First Amendment content neutrality).

128. *See* Marc. S. Hanish, *California Supreme Court Survey—Bail Conditions*, 23 PEPP. L. REV. 286, 289–90 (1996) (noting three kinds of cases where judicial reformation might occur: "(1) cases involving procedural safeguards required by the First Amendment or due process; (2) cases involving classifications that were held to be underinclusive under the equal protection clause; and (3) cases involving otherwise vague or overbroad criminal statutes."); *see also* N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50 (1982) (invalidating a statute creating bankruptcy judges, in part because they lack life tenure).

129. *See* Welsh v. United States, 398 U.S. 333, 366 n.18 (1970) (Harlan, J., concurring in result); Ginsburg, *supra* note 22, at 324.

leave and men receive zero, judges can rewrite the law to give men ten.[130] Similarly, if a public benefits law gives men $5,000 a year and women $2,000, equality rights can be protected by leveling women up to $5,000 (or men down to $2,000). Judges thus do not face the problem that they see in other kinds of cases, like vagueness or due process cases, where they may have to exercise more creativity in rewriting the statute. Second, Purpose Preservation for equality cases first emerged in suits challenging laws that provided social security and similar welfare benefits, and invalidating these laws would have harmed some vulnerable groups. Extension is appealing in such cases (just as in *Liberta*, where invalidation would have legalized rape).[131] This concern might have been especially pronounced in cases like *Weinberger*, *Frontiero*, etc., where sex equality claims were brought against statutes granting women greater benefits than men. It does seem ironic to eliminate a benefit designed for women in the name of gender equality.[132] Thus the different treatment of equality cases may be a product of historical contingency. None of these concerns, however, are reflected in the official doctrine, which simply instructs that the legislative purpose determines whether a statute is leveled up or down after an equality challenge.[133]

B. Severability

Severability doctrine concerns situations where a court has determined that one provision of a statute is unconstitutional, and is thereby empowered to invalidate other provisions as well.[134] Classically, American severability doctrine is

130. One interesting question is whether judges can "level middle," that is, whether they can level down the included class and level up the excluded class such that they meet somewhere in the middle. In *Schachter v. Canada*, for instance, adoptive parents had been given 15 weeks of maternity leave, and natural parents zero. The legislature ultimately amended the law after the case was brought to give both groups ten weeks. [1992] 2 S.C.R. 679, 724 (Can.). Could the court have imposed this solution itself? It seems unlikely that the legislative history would point judges to a specific intermediate number, and so in practice they focus on the numbers in the statute (e.g. fifteen versus zero). But in some cases it might be clear that the legislature would not want to level up all the way or down all the way, for instance if leveling up would be extremely expensive and leveling down would deprive a group of a much needed benefit. In such cases, leveling middle would make sense even in the face of ambiguity about which number to choose.

131. One scholar even called for extension to be the remedy in all cases. *See* Kovacic, *supra* note 123, at 45. Another argued that extension should be the remedy where the government program provides "food, shelter, and other necessities of life." Beers, *supra* note 125, at 139 (quoting Shapiro v. Thompson, 394 U.S. 618, 627 (1969)).

132. *See* Caminker, *supra* note 11, at 1196–1202; *supra* note 84 and accompanying text.

133. *See* Heckler v. Mathews, 465 U.S. 728, 739 (1984).

134. Severability determinations differ from the other remedial questions discussed in this Part because they do not involve two different ways of solving a constitutional problem. Rather, they involve judges determining what else to strike down once a provision has been invalidated. Whether such determinations are considered discretionary or not will depend on one's theory of why judges can

governed by the logic of Purpose Preservation.[135] When a judge finds a statutory provision unconstitutional, they apply two tests: first, "whether the statute will function in a *manner* consistent with the intent of Congress," and second, whether "the statute created in [the unconstitutional provision's] absence is legislation that Congress would not have enacted."[136] Each of these tests gets at the question of whether declaring further provisions inseverable will advance the legislative purpose. If knocking out the unconstitutional provision will so warp the rest of the statute that it no longer fits with Congress's intentions, or that Congress would not have enacted it, then the rest of the statute must fall. By contrast, an Editorial Restraint approach to severability would deny judges the power to decide what stays and what goes according to their reading of the legislative will. This could be achieved through an approach that treats all provisions as severable, such that when a provision is struck down judges cannot then invalidate further provisions without the legislature's instruction.[137] Or it could be achieved by making no provisions severable, such that when a provision is struck down the entire statute must fall.[138] But federal doctrine has adopted neither of these Editorial Restraint strategies, and instead embraces the logic of Purpose Preservation.

However, there is an odd disconnect between doctrine and practice in the Supreme Court's severability decisions. Officially, whenever the Court strikes down a provision it can rummage through the rest of the statute to see what else should go, according to its understanding of the legislature's purposes. But in practice, the modern Supreme Court very rarely declares statutory provisions inseverable. Prior to 1940 the Court did so with some frequency, but since that time it has found statutes inseverable in only one case involving a federal law and only a handful more involving state laws.[139] The Court has at times been criticized for this reticence, for example in legislative veto cases, with commentators arguing that the severed statutes would not have been enacted without legislative

declare provisions inseverable in the first place. *See* Fish, *supra* note 23, at 1323. For further discussion, see *infra* Part V.C.

135. For a fuller exposition of the history of severability doctrine, see Fish, *supra* note 23, at 1300–09; Kenneth A. Klukowski, *Severability Doctrine: How Much of a Statute Should Federal Courts Invalidate?*, 16 TEX. REV. L & POL. 1, 10–30 (2011).

136. Alaska Airlines, Inc. v. Brock, 480 U.S. 678, 685 (1987); *accord* Immigration & Naturalization Serv. v. Chadha, 462 U.S. 919, 931–32 (1983); Champlin Ref. Co. v. Corp. Comm'n of Okla., 286 U.S. 210, 234 (1932). There is also a third test, looking at whether the remaining law can operate independently of the severed portion. But this test is rather narrow, and has been deemphasized in recent cases. *See, e.g.*, *Alaska Airlines*, 480 U.S. at 684; Fish, *supra* note 23, at 1332.

137. For authors advocating such an approach, see Tobias A. Dorsey, Remark, *Sense and Severability*, 46 U. RICH. L. REV. 877 (2012); Fish, *supra* note 23; Walsh, *supra* note 8.

138. *See* Campbell, *supra* note 54.

139. *See* Fish, *supra* note 23, at 1303.

veto provisions.[140] And the Court has also in recent decades adopted a default presumption of severability, such that statutory provisions are assumed to be severable unless there is clear evidence to the contrary, reversing the Court's prior default presumption of inseverability.[141] Whether due to this presumption, a general reticence to mark its red pen over valid statutory provisions, or both, the modern Supreme Court finds very few laws inseverable. Thus, while the official severability doctrine reads like a textbook case of Purpose Preservation, in practice the Court seems to have embraced Editorial Restraint.

C. Facial or As-Applied Invalidation

When the Supreme Court decides whether to strike down a provision facially or as applied, it is deciding whether to invalidate the entire provision or merely to declare that it does not apply in a certain kind of case. The Court's official doctrine on this question follows the logic of Editorial Restraint—it establishes a strict preference for as-applied challenges, thus preventing courts from exercising creative judgment in deciding whether to strike down the entire statute or just one application. As the Court announced in *United States v. Salerno*, in facial challenges "the challenger must establish that no set of circumstances exists under which the Act would be valid."[142] Thus courts engaged in judicial review can only create an as-applied exception to the relevant statute, unless the statute has no possible valid application. The Supreme Court has justified this strong disfavor for facial challenges by arguing that they "run contrary to the fundamental principle of judicial restraint that courts should neither 'anticipate a question of constitutional law in advance of the necessity of deciding it' nor 'formulate a rule of constitutional law broader than is required by the precise facts to which it is to be applied.'"[143] Such language powerfully echoes the themes of Editorial Restraint: judges should stick to the facts before them, and do no more than necessary to decide the case.

Courts considering facial invalidation thus do not, as they do in severability decisions, determine which remedy better serves the legislative purpose. This difference of approach seems a bit strange, given how similar severability is to the

140. *See, e.g.*, Campbell, *supra* note 54, at 1523–24; Glenn Chatmas Smith, *From Unnecessary Surgery to Plastic Surgery: A New Approach to the Legislative Veto Severability Cases*, 24 HARV. J. ON LEGIS. 397, 407–19 (1987); Tribe, *supra* note 8, at 22.

141. *Compare* Regan v. Time, Inc., 468 U.S. 641, 653 (1984), *with* Williams v. Standard Oil Co., 278 U.S. 235, 242 (1929).

142. 481 U.S. 739, 745 (1987).

143. Wash. State Grange v. Wash. State Republican Party, 552 U.S. 442, 450 (2008) (quoting Ashwander v. Tenn. Valley Auth., 297 U.S. 288, 346–47 (1936) (Brandeis, J., concurring)).

distinction between as-applied and facial challenges. Both kinds of cases involve a court deciding, after a constitutional violation has been found, whether to eliminate further parts of the statute—in severability cases by striking down otherwise valid provisions, and in facial challenge cases by striking down otherwise valid (or potentially valid) applications. Indeed, many scholars have connected severability and facial challenges, treating them as involving the same basic questions.[144] And the Supreme Court itself has at times done the same, discussing severability of provisions and applications interchangeably.[145] This essential similarity makes the doctrinal focus on Purpose Preservation for severability questions and Editorial Restraint for facial challenges difficult to understand. If anything, the approaches should be reversed. The argument for Editorial Restraint is stronger for severability doctrine, given that it empowers judges to search anywhere else in the statute (or perhaps even the legislative code) for inseverable provisions, while facial challenges limit judges to only the provision being litigated.[146]

The official doctrine does, however, allow some exceptions to this preference for as-applied challenges. One such exception is First Amendment overbreadth cases. If a statute prohibits some constitutionally protected speech, and also prohibits some unprotected speech, the Court will strike the statute down on its face. This is because an overbroad statute prohibiting speech can chill protected speech, as potential speakers will not know where the constitutional line is drawn in advance of litigation.[147] Another exception exists for statutes enacted for an unconstitutional purpose. Such statutes can be struck down facially, even if they have otherwise valid applications, because of the legislature's unconstitutional intent.[148] For example, in *Edwards v. Aguillard*, the Supreme Court struck down a Louisiana law providing that creationism be taught in schools, and did so facially because

144. *See, e.g.*, RICHARD H. FALLON, JR. ET AL., HART AND WECHSLER'S THE FEDERAL COURTS AND THE FEDERAL SYSTEM 182 (5th ed. 2003); Dorf, *supra* note 110, at 249–51; Gillian E. Metzger, *Facial Challenges and Federalism*, 105 COLUM. L. REV. 873, 886 (2005); Robert L. Stern, *Separability and Separability Clauses in the Supreme Court*, 51 HARV. L. REV. 76, 78–79 (1937); Adrian Vermeule, *Saving Constructions*, 85 GEO. L.J. 1945, 1950 n.26 (1997).

145. *See, e.g.*, Ayotte v. Planned Parenthood, 546 U.S. 320, 328–29 (2006); United States v. Booker, 543 U.S. 220, 320 (2005) (Thomas, J., dissenting in part).

146. *See* Fish, *supra* note 23, at 1316–17 (discussing several possible units of analysis for severability: the bill, the act, and the code).

147. *See* Virginia v. Hicks, 539 U.S. 113, 119 (2003). As some scholars have pointed out, the rationale for overbreadth doctrine is not limited to the First Amendment context, although the Court has yet to explicitly expand it beyond that context. *See* Dorf, *supra* note 110, at 264–68; Richard H. Fallon, Jr., *Making Sense of Overbreadth*, 100 YALE L.J. 853, 862, 884 n.192 (1991); *see also Salerno*, 481 U.S. at 745 ("[W]e have not recognized an 'overbreadth' doctrine outside the limited context of the First Amendment"); Schall v. Martin, 467 U.S. 253, 268–69 n.18 (1984).

148. *See* Dorf, *supra* note 110, at 279–80.

the purpose of the statute was to advance religion.[149] Both of these exceptions can be understood by reference to the concept of Irony Avoidance discussed above.[150] In both kinds of cases, to follow the normal rule and entertain only as-applied challenges would produce a constitutional irony. In overbreadth cases, it would leave people afraid to exercise the very rights being protected, and in unconstitutional purpose cases, it would preserve a statute intended to violate the constitutional rule being enforced. These exceptions can thus be understood as limited allowances for constitutional norms in a doctrinal scheme that generally emphasizes Editorial Restraint.

However, the Supreme Court's actual practice concerning facial challenges does not fit neatly with the official doctrine. While *Salerno* seems to establish a strict preference for as-applied challenges, the Court in fact frequently entertains facial challenges even outside of the two circumstances just discussed. This fact has been observed by a number of scholarly commentators, most notably Richard Fallon, who has shown that the Supreme Court regularly entertains facial challenges.[151] The disconnect between doctrine and practice has even been noted by members of the Court itself. Justices Stevens, Ginsburg, and Scalia have each taken their colleagues to task, in various circumstances, for failing to adhere to the stated preference for as-applied challenges.[152] This is an interesting reversal of the pattern seen in the Court's severability doctrine. In the severability context the Court purports to adopt Purpose Preservation but in practice almost never finds statutes inseverable, and in the facial challenges context it purports to adopt Editorial Restraint but in practice commonly permits facial challenges. However, as Fallon has observed, the choice between as-applied and facial challenges is

149. 482 U.S. 578 (1987).

150. *See supra* Part II.C.

151. Richard H. Fallon, Jr., *Fact and Fiction About Facial Challenges*, 99 CALIF. L. REV. 915, 918 (2011) ("In all of those Terms, the Court adjudicated more facial challenges on the merits than it did as-applied challenges."); *see also* Dorf, *supra* note 110, at 236, 238; Fallon, *supra* note 53, at 1323; Marc E. Isserles, *Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement*, 48 AM. U. L. REV. 359, 439 (1998).

152. *See, e.g.*, Shelby Cty. v. Holder, 133 S. Ct. 2612, 2648 (2013) (Ginsburg, J., dissenting); Citizens United v. Fed. Election Comm'n, 558 U.S. 310, 398–405 (2010) (Stevens, J., concurring in part and dissenting in part) (arguing that the case should be decided as an as-applied challenge); Gonzales v. Carhart, 550 U.S. 124, 187 (2007) (Ginsburg, J., dissenting); Chicago v. Morales, 527 U.S. 41, 81 (1999) (Scalia, J., dissenting); Janklow v. Planned Parenthood, 517 U.S. 1174, 1175 (1996) (Stevens, J., respecting denial of petition for certiorari) ("While a facial challenge may be more difficult to mount than an as-applied challenge, the dicta in *Salerno* 'does not accurately characterize the standard for deciding facial challenges,' and 'neither accurately reflects the Court's practice with respect to facial challenges, nor is it consistent with a wide array of legal principles.'") (quoting Dorf, *supra* note 110, at 236, 238); Ada v. Guam Soc'y of Obstetricians & Gynecologists, 506 U.S. 1011 (1992) (Scalia, J., dissenting) (arguing that facial challenges to laws restricting abortion should be decided under the *Salerno* "no set of circumstances" test).

driven largely by case-specific doctrinal issues, and not by an overarching concern for the statutory purpose.[153]

One could nonetheless imagine a facial challenge doctrine following the logic of Purpose Preservation—much like severability doctrine, facial challenge doctrine could instruct courts to ask whether the law still furthers the legislature's intended purposes with certain applications removed. Indeed, a number of the Court's decisions hint at just such an approach.[154] In *Northern Pipeline*, the Supreme Court decided that bankruptcy courts could not decide state law contract claims.[155] However it also removed bankruptcy courts' jurisdiction over other kinds of cases, since it concluded that Congress would not have preserved those other areas of jurisdiction had it known that jurisdiction over contract claims was unconstitutional.[156] Similarly, in *Butts v. Merchants and Miners Transportation Co.*, the Court had to decide whether the Civil Rights Act of 1875 applied in the federal territories and on American ships, when it had previously been struck down as applied to activity within the states.[157] The Court concluded that Congress would not have passed the Civil Rights Act had it known the Act would only be applied to the territories and to ships.[158] These cases seem to establish the principle that Congress would not have wanted an idiosyncratic remainder—if a statute is struck down as to the vast majority of applications, the court should not keep it in place for the few remaining because doing so would run contrary to Congress's purposes. However the Court has not explicitly stated this idea as doctrine, nor, more broadly, has it embraced Purpose Preservation as a governing theory for facial challenges in any sort of general, prospective statement. *Salerno* remains the official rule, and the Court does not appear to acknowledge the conflict between *Salerno* and the cases that use Purpose Preservation-based analysis.

D. Constitutional Avoidance

The Supreme Court has long held that if there are two different ways to interpret a statute, one that creates constitutional violation and one that does not, then a judge should adopt the interpretation that renders the statute

153. *See generally* Fallon, *supra* note 151; Fallon, *supra* note 53; *accord* Dorf, *supra* note 110.
154. *See, e.g.*, Ayotte v. Planned Parenthood, 546 U.S. 320, 330 (2006); United States v. Nat'l Treasury Emps. Union, 513 U.S. 454, 479 (1995); Sloan v. Lemon, 413 U.S. 825, 834 (1973).
155. N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 87 n.40 (1982).
156. *Id.*
157. 230 U.S. 126 (1913).
158. *Id.* at 138.

constitutional.[159] This doctrine of constitutional avoidance has become canonical, to the point that the Court has stated that it is "beyond debate."[160] And, much as in England, the avoidance canon is mandatory on American courts: "every reasonable construction must be resorted to, in order to save a statute from unconstitutionality."[161] The classic formulation of the avoidance canon does not seem to present a conflict between Editorial Restraint and Purpose Preservation. Indeed, it assumes away any such conflict—if both possible interpretations are equally plausible, then choosing the constitutional interpretation cannot conflict with the legislative purpose.[162] Further, constitutional avoidance is formally a merits issue, not a remedial issue, since it determines the meaning of the statute rather than changing the statute to make it constitutional. While avoidance does the same work as a constitutional remedy, it is understood to be simply a tool of interpretation.[163]

However, in practice the avoidance canon does not merely function as an interpretive tiebreaker. In a number of Supreme Court cases, it is instead a way of effectively changing the statute's meaning.[164] A few recent high-profile cases illustrate this usage. In *Northwest Austin Municipal Utility District Number One v. Holder*, the Court interpreted the phrase "political subdivision" in Section V of the Voting Rights Act to include tiny utility districts, even though the statute

159. *See, e.g.*, Ashwander v. Tenn. Valley Auth., 297 U.S. 288, 348 (1936) (Brandeis, J., concurring). Modern uses of the avoidance canon have done away with the requirement of an actual violation, holding that "constitutional doubts" are sufficient to prompt an avoidance reading. *See* United States v. Del. & Hudson Co., 213 U.S. 366, 389 (1909); Neal Kumar Katyal & Thomas P. Schmidt, *Active Avoidance: The Modern Supreme Court and Legal Change*, 128 HARV. L. REV. 2109, 2117 (2015); Vermeule, *supra* note 144, at 1949.

160. Edward J. DeBartolo Corp. v. Fla. Gulf Coast Bldg. & Constr. Trades Council, 485 U.S. 568, 575 (1988); *see also* Spector Motor Serv. v. McLaughlin, 323 U.S. 101, 105 (1944).

161. *DeBartolo*, 485 U.S. at 575 (quoting Hooper v. California, 155 U.S. 648, 657 (1895)); *see also* Virginia v. Am. Booksellers Ass'n, 484 U.S. 383, 397 (1988); United States v. Batchelder, 442 U.S. 114, 122 (1979).

162. Indeed, the Court has even sometimes suggested that the avoidance canon is a rule of thumb to determine legislative intent. *See, e.g.*, Clark v. Martinez, 543 U.S. 371, 382 (2005) ("The canon is thus a means of giving effect to congressional intent, not of subverting it."); Rust v. Sullivan, 500 U.S. 173, 191 (1991) ("This canon is followed out of respect for Congress, which we assume legislates in the light of constitutional limitations.").

163. *But see* United States v. Booker, 543 U.S. 220, 249 (2005) (referring to the dissenters' proposed avoidance reading as a "remedy"); *id.* at 292 (Stevens, J., dissenting in part) (same); United States v. Marshall, 908 F.2d 1312, 1318 (7th Cir. 1990) ("The canon about avoiding constitutional decisions, in particular, must be used with care, for it is a closer cousin to invalidation than to interpretation. It is a way to enforce the constitutional penumbra, and therefore an aspect of constitutional law proper."); ROACH, *supra* note 60, at 14 ("Although interpreting a statute is not technically a remedy for a constitutional violation, it is, in a functional sense, an alternative to invalidation of a potentially unconstitutional law.").

164. *See* Katyal & Schmidt, *supra* note 159.

explicitly defined that term in a way that excluded such entities.[165] Commentators have pointed out that this interpretation was highly implausible.[166] But the Court adopted it nonetheless, so as to avoid deciding the constitutionality of the Voting Rights Act.[167] In a subsequent case, *National Federation of Independent Business v. Sebelius*, the Court was even more transparent.[168] In that case, the Court interpreted a provision that fined people for not purchasing health insurance as a tax, rather than a command, making it constitutionally permissible under Congress's taxing power.[169] However, Chief Justice Roberts's majority opinion explicitly noted that the provision was most naturally read as a command, and that it could only be interpreted as a tax because as a command it would be unconstitutional.[170] Thus the Court acknowledged that it was, under the guise of interpretation, changing the meaning of the statute to avoid finding it unconstitutional. Such cases are not a new development—the Court has been using the avoidance canon to rewrite statutes for some time, even if it had not previously announced that it was doing so.[171] Viewed in this light, the categorical rule favoring avoidance readings takes on the logic of Editorial Restraint—much as in England, American courts avoid invalidating laws whenever possible by stretching their meaning to avoid a constitutional problem.[172] Even if an interpretation

165. 557 U.S. 193 (2009).

166. *See, e.g.*, Jonathan H. Adler, *Judicial Minimalism, the Mandate, and Mr. Roberts*, *in* THE HEALTH CARE CASE: THE SUPREME COURT'S DECISION AND ITS IMPLICATIONS 174 (Nathaniel Persily et al. eds., 2013); Travis Crum, Note, *The Voting Rights Act's Secret Weapon: Pocket Trigger Litigation and Dynamic Preclearance*, 119 YALE L.J. 1992, 1996 (2010); Richard L. Hasen, *Constitutional Avoidance and Anti-Avoidance by the Roberts Court*, 2009 SUP. CT. REV. 181 (2009); Katyal & Schmidt, *supra* note 159, at 2129–34; Heather K. Gerken, *Online VRA symposium: Reading the Tea Leaves—the Uncertain Future of the Act*, SCOTUSBLOG (Sept. 11, 2012, 1:40 PM), http://www.scotusblog.com/2012/09/online-vra-symposium-reading-the-tea-leaves-the-uncertain-future-of-the-act/ [https://perma.cc/A3JQ-4QAW].

167. The Court nonetheless reached the question in a subsequent case, *Shelby County v. Holder*, 133 S. Ct. 2612 (2013).

168. 132 S. Ct. 2566 (2012).

169. *Id.* at 2600.

170. *Id.* at 2600–01 ("[T]he statute reads more naturally as a command to buy insurance than as a tax, and I would uphold it as a command if the Constitution allowed it. . . . [I]t is only because we have a duty to construe a statute to save it, if fairly possible, that § 5000A can be interpreted as a tax.").

171. *See, e.g.*, Philip P. Frickey, *Getting From Joe to Gene (McCarthy): The Avoidance Canon, Legal Process Theory, and Narrowing Statutory Interpretation in the Early Warren Court*, 93 CALIF. L. REV. 397, 402, 460–61 (2005); Frederick Schauer, Ashwander *Revisited*, 1995 SUP. CT. REV. 71, 74–80 (1995). The Court's Justices have also commonly pointed out when their colleagues make liberal use of the avoidance canon. *See, e.g.*, Bond v. United States, 134 S. Ct. 2077, 2095–97 (2014) (Scalia, J., concurring); Skilling v. United States, 130 S. Ct. 2896, 2939–41 (2010) (Scalia, J., concurring in part and concurring in the judgment); United States v. X-Citement Video, Inc., 513 U.S. 64, 80 (1994) (Scalia, J., dissenting); Welsh v. United States, 398 U.S. 333, 345 (1970) (Harlan, J., concurring in result); Int'l Ass'n of Machinists v. Street, 367 U.S. 740, 785–86 (1961) (Black, J., dissenting).

172. *See supra* Part II.A.

runs contrary to what the legislature intended, it will be adopted so long as it is sufficiently plausible and elides the constitutional issue.[173] Courts thereby avoid taking on the further degree of editorial power required to strike down the statute, and stick to the essentially judicial practice of interpreting laws, rather than the more legislative practice of deleting them.

If constitutional avoidance is treated not as a tiebreaker between plausible interpretations, but as a way of *re*interpreting the statute to make it constitutional, that also opens up the possibility of a Purpose Preservation approach to avoidance. Rather than always adopting the avoidance reading, as the mandatory language in avoidance cases seems to require,[174] a reviewing court could consider it alongside possible remedies, such as invalidation, and adopt the fix that best furthers the purposes of the enacting legislature. Indeed, the Supreme Court did precisely this in *United States v. Booker*.[175] In *Booker*, the Court found that the Sentencing Reform Act violated the Sixth Amendment right to a trial by jury. The Act created a system of sentencing guidelines that made certain judge-found facts lead to a mandatory increase in the defendant's sentence, without requiring that those facts be proven to a jury. After this violation was established, however, the justices disagreed on the proper remedy. Justice Stevens wanted to adopt an avoidance interpretation, construing the word "court" in the Act to mean "judge and jury," and not just "judge."[176] This would have meant that any guidelines enhancements must be proven to a jury beyond a reasonable doubt, as any other element of a crime must be. Justice Breyer instead argued that certain provisions of the Act should be struck down and others found inseverable in order to make the Sentencing Guidelines advisory, so that enhancement factors would no longer cause a mandatory increase in sentence.[177] Each Justice argued for his respective solution by trying to show that it better fit with the enacting Congress's goals in passing the Sentencing Reform Act, and that the alternative solution would contradict the choices Congress had made.[178] Justice Breyer ultimately wrote the majority opinion on the remedial question. The Court thus determined that Justice Stevens's avoidance remedy would deviate further from Congress's intentions than would striking down portions of the

173. *See* Vermeule, *supra* note 144 (arguing that the avoidance canon overprotects constitutional norms at the expense of legislative intent).

174. *See* cases cited *supra* note 161.

175. 543 U.S. 220 (2005).

176. *Id.* at 286 (Stevens, J., dissenting in part) ("As a textual matter, the word 'court' can certainly be read to include a judge's selection of a sentence as supported by a jury verdict—this reading is plausible either as a pure matter of statutory construction or under principles of constitutional avoidance.").

177. *Id.* at 258–65 (majority opinion).

178. *Id.* at 247–49; *id.* at 292 (Stevens, J., dissenting in part).

statute, and so it opted for partial invalidation. In sum, the Court in *Booker* adopted the logic of Purpose Preservation to reject an avoidance reading of a statute.

E. Soliciting a Legislative Remedy

A final remedial issue to consider is whether and when the Supreme Court solicits a remedy from the enacting legislature, rather than imposing a remedy itself. Seeking a legislative remedy could be seen as consistent with both Editorial Restraint and Purpose Preservation, since it means that the Court does not have to change the statute, and it also means that the legislature can select the remedy that best furthers its own goals.[179] England and Canada both have established mechanisms to do this. In England it is the last option if no avoidance interpretation is available—the reviewing court can only make a nonbinding "declaration of incompatibility," and thereby ask the legislature to amend the rights-infringing statute.[180] In Canada, courts have the power to invalidate or otherwise fix unconstitutional statutes, but they sometimes delay their remedies so as to grant the legislature time to craft its own. In *Schachter*, the Canadian Supreme Court stated that such delay is permissible if the remedy is likely to cause serious harm that a legislative fix would help alleviate.[181] In the United States, by contrast, the Supreme Court virtually never solicits a legislative remedy to a constitutional violation or delays its own remedy to give the legislature a chance to act. While the Court does sometimes remand remedial issues to state courts,[182] it rarely punts them to state legislatures, much less to Congress. The most notable attempt to explicitly solicit a legislative remedy was in *Northern Pipeline Construction Co. v. Marathon Pipe Line Co.*,[183] where the Court delayed its remedy for several months to give Congress the chance to pass a new law.[184] This attempt failed badly.

179. Of course, a remedy can only be solicited from the contemporaneous legislature, which will usually have a different composition from the legislature that enacted the statute. For more on why this distinction matters, see *infra* Part V.D.

180. *See supra* notes 40–46 and accompanying text.

181. Schachter v. Canada, [1992] 2 S.C.R. 679, 719 (Can.). South Africa has a similar approach. *See* Pius N. Langa, *The Role of the Constitutional Court in the Enforcement and Protection of Human Rights in South Africa*, 41 ST. LOUIS U. L.J. 1259, 1274–75 (1997). Indeed, the South African Constitution explicitly permits judges to delay remedies. *See* S. AFR. CONST., 1996 § 172 (permitting "an order suspending the declaration of invalidity for any period and on any conditions, to allow the competent authority to correct the defect").

182. *See infra* Part V.B.

183. 458 U.S. 50, 87–89 (1982).

184. The Court also delayed the remedy in *Buckley v. Valeo* by thirty days. 424 U.S. 1, 144 (1976). However, that delay does not seem to have been intended to solicit a legislative fix, but rather to give the Federal Election Commission time to exercise its powers.

In *Northern Pipeline*, the Supreme Court held that the system of bankruptcy courts set up by the Bankruptcy Act of 1978 was unconstitutional.[185] The Court determined that bankruptcy judges had been granted Article III powers without having also been granted the protections of Article III, such as life tenure and immunity from diminution of salary.[186] Recognizing how vast the reliance interests were in this case, the Court decided to make its remedy non-retroactive, and to stay its judgment for roughly three months so as to "afford Congress an opportunity to reconstitute the bankruptcy courts or to adopt other valid means of adjudication, without impairing the interim administration of the bankruptcy laws."[187] When this stay had nearly ended the Solicitor General moved for an extension, and the Court granted one for roughly three more months.[188] When the extension had nearly ended the Solicitor General petitioned for another, which was denied.[189] Congress had, meanwhile, not replaced the invalidated system of bankruptcy courts, as the House and Senate had deadlocked over conflicting solutions.[190] And so instead, on Christmas 1982 (the day after the Supreme Court's judgment went into effect), the Judicial Conference of the United States established an "Emergency Rule" which it urged the federal circuits to adopt.[191] The Emergency Rule made bankruptcy judges the equivalent of special masters—federal district judges could refer bankruptcy cases to them, but all of their findings would be reviewed *de novo*. Thus the judiciary itself crafted a temporary solution, though many at the time believed that doing so was beyond its powers.[192] The Emergency Rule was challenged in the federal courts, and several courts of appeals upheld it, though the Supreme Court never granted certiorari on the issue.[193] Finally, Congress enacted a legislative solution nearly two years after

185. *Northern Pipeline*, 458 U.S. at 87.
186. *Id.* at 61–62.
187. *Id.* at 88.
188. N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 459 U.S. 813 (1982) (granting stay until December 24, 1982).
189. N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 459 U.S. 1094 (1982) (denying stay).
190. *See* Tamar Lewin, *Confusion Over Status of Bankruptcy Court*, N.Y. TIMES (Dec. 28, 1982), http://www.nytimes.com/1982/12/28/business/confusion-over-status-of-bankruptcy-court.html [https://perma.cc/GUM7-5W8K] ("That question has become a political one, however, with the Senate, the House of Representatives and the United States Judicial Conference each committed to a different answer.").
191. *See* Lawrence P. King, *The Unmaking of a Bankruptcy Court: Aftermath of* Northern Pipeline v. Marathon, 40 WASH. & LEE L. REV. 99, 115–16 (1983).
192. *See id.* at 116–17; Lewin, *supra* note 190 ("The rules will be in effect either until Congress finds a way to deal with the Supreme Court's June 28 decision that the Federal bankruptcy law is unconstitutional or until the interim rules are deemed illegal, as many constitutional lawyers believe they will be.").
193. *See* Salomon v. Kaiser, 722 F.2d 1574 (2d Cir. 1983); Coastal Steel Corp. v. Tilghman Wheelabrator Ltd., 709 F.2d 190 (3d Cir. 1983), *cert. denied*, 464 U.S. 938 (1983); White Motor

the decision in *Northern Pipeline*—the Bankruptcy Amendments and Federal Judgeship Act of 1984.[194] This episode illustrates how difficult it is for judges to compel Congress to act within a given timeline even where the consequences of inaction are potentially catastrophic.[195] It may also help explain why the Supreme Court has not tried to solicit a legislative solution to a constitutional violation since *Northern Pipeline*.

But the Court has on occasion fired warning shots, signaling to the elected branches that it intends to find a law unconstitutional sometime in the near future unless there is a legislative fix. Richard Re refers to this strategy as "the doctrine of one last chance."[196] The Court prominently employed it in *Northwest Austin Municipal Utility District No. 1 v. Holder (NAMUDNO)*, where commentators viewed the adoption of an implausible avoidance interpretation—joined by liberal as well as conservative members of the Court—as signaling a strong likelihood that the Voting Rights Act would soon be struck down unless Congress enacted a version more tailored current voting suppression patterns.[197] This ex ante solicitation of a legislative remedy did not prompt legislative action, and the Court carried through on its implicit threat.[198] Re observes that the Court adopted a similar strategy in other major cases, perhaps most notably the detention cases—*Rasul v. Bush*, *Hamdan v. Rumsfeld*, and *Boumediene v. Bush*—in which the Court's decisions repeatedly prompted congressional action, which was then reviewed anew by the Court in a dialogic fashion.[199] In a way, this ex ante signaling is a bit more respectful of the legislature than the *Northern Pipeline* approach. Rather than announcing its ruling as a time bomb and giving Congress a countdown to defuse it, the Court lets Congress know ahead of time that there are likely constitutional issues. It

Corp. v. Citibank, N.A., 704 F.2d 254, 262 (6th Cir. 1983); First Nat'l Bank of Tekamah v. Hansen, 702 F.2d 728 (8th Cir. 1983) (per curiam), *cert. denied*, 463 U.S. 1208 (1983); Braniff Airways, Inc. v. Civil Aeronautics Bd., 700 F.2d 214 (5th Cir. 1983), *cert. denied*, 461 U.S. 944 (1983); Stewart v. Stewart (*In re* Stewart), 741 F.2d 127, 130 (7th Cir. 1984); Okla. Health Servs. Fed. Credit Union v. Webb, 726 F.2d 624 (10th Cir. 1984); Lucas v. Thomas (*In re* Thomas), 765 F.2d 926, 929 (9th Cir. 1985).

194. Pub. L. No. 98-353, 98 Stat. 333, 345 (1984).

195. *See also* Matthew R. Christiansen & William N. Eskridge, Jr., *Congressional Overrides of Supreme Court Statutory Interpretation Decisions, 1967–2011*, 92 TEX. L. REV. 1317 (2014) (showing that congressional overrides of Supreme Court statutory decisions have, in the last decade or so, fallen off dramatically).

196. Richard M. Re, *The Doctrine of One Last Chance*, 17 GREEN BAG 2D 173 (2014).

197. 557 U.S. 193 (2009); *see also* Crum, *supra* note 166, at 1996; Re, *supra* note 196, at 175.

198. *See* Shelby Cty. v. Holder, 133 S. Ct. 2612 (2013).

199. Boumediene v. Bush, 553 U.S. 723 (2008); Hamdan v. Rumsfeld, 548 U.S. 557 (2006); Rasul v. Bush, 542 U.S. 466 (2004); *see also* Re, *supra* note 196, at 177.

thereby gives Congress time, fair warning, and an opportunity to fix the problem, without creating a dramatic crisis.

IV. WHICH APPROACH IS BETTER?

The first three Parts of this article have been a mapping exercise. They have shown that the selection of constitutional remedies is a matter of judicial discretion, that this discretion can be exercised through the logic of Editorial Restraint or Purpose Preservation, and that the American doctrine of constitutional remedies is a confounding mishmash of both approaches. This Part argues that we should follow Canada's lead, and reconstruct American doctrine according to the logic of Purpose Preservation. I begin by critiquing Editorial Restraint, showing that while it works in a system where judges lack the power to change laws, such as England's, it makes little sense in a system with strong-form judicial review. I then mount a defense of Purpose Preservation, arguing that it furthers the value of democratic governance.

A. Editorial Restraint is Mostly Empty Formalism

Editorial Restraint is animated by the intuition that legislatures, not judges, are supposed to change statutes. Thus judges in England cannot invalidate statutes that violate protected rights—they can only ask the legislature to do so. This intuition carries over into other remedial choices as well. Adding language to a statute seems more legislative than striking the statute down, striking it down seems more legislative than creating an as-applied exception, and creating such an exception seems more legislative than interpreting the statute to avoid the constitutional issue. With each step down this remedial ladder, the action that the court takes appears to give the court a smaller amount of power to change the statute. And the less editorial power judges assume, the better. However this intuition does not withstand scrutiny. Facial invalidation, rewriting, as-applied exceptions, and avoidance interpretations all involve courts taking the same basic action—changing the statute's meaning in order to fix a constitutional violation. Indeed, these remedies are quite difficult to distinguish in practice. The notion that any one of them requires courts to take on more legislative power than the others is, for the most part, empty formalism.

There is, however, one exception. When judges solicit a remedy from the relevant legislature, they do exercise less editorial control over legislation than if they simply changed the statute themselves. Thus England's reliance on

declarations of incompatibility draws a coherent line between what is legislative and what is judicial.[200] If judges cannot change statutes at all, but can only request a legislative fix when they find a law unconstitutional, then the legislature's monopoly on editorial control over statutes is preserved. But once we move away from the English system and towards a system that grants judges the power to change statutes, the logic of Editorial Restraint breaks down. So the United States must either be like England and reject strong-form constitutional review altogether, or abandon the principle of Editorial Restraint.[201] There is no middle ground.

Consider first the distinction between adding language to a statute and striking language down. Intuitively, the former seems like a worse intrusion into the legislative sphere. For example, in *People v. Liberta* the New York Court of Appeals effectively rewrote the New York rape statute, changing it from "A male is guilty of rape in the first degree when he engages in sexual intercourse with a female . . . by forcible compulsion" to "A male [or female] is guilty of rape in the first degree when he [or she] engages in sexual intercourse with a female [or male] . . . by forcible compulsion."[202] Is there any justification for the intuition that adding words like this is a greater assumption of legislative power than striking the statute down? Both possible remedies change the statute, and it is difficult to articulate why a worse change is wrought by adding language than by deleting language. This difficulty is compounded by the fact that the difference between adding and deleting language is determined by arbitrary choices in legislative drafting. Had the statute been written in sex-neutral terms, and the phrase "except that a female cannot be found guilty of rape, and a male cannot be a victim of rape" been tacked on at the end, then the New York Court of Appeals could have achieved the exact same result by invalidating that phrase. But why is striking down a textual exception better than striking down an atextual exception? As the Canadian Supreme Court observed in *Schachter*, "It would be an arbitrary distinction to treat inclusively and exclusively worded statutes differently. To do so would create a situation where the style of drafting would be the single critical factor in the determination of a remedy."[203] Adding language to a statute is just as much of an intrusion into the legislative domain as deleting language—both involve changing the statute's

200. *See supra* notes 40–46 and accompanying text.

201. Note, however, that even England does not have a system that keeps editorial control purely in the hands of the legislature. The Human Rights Act also permits (indeed requires) judges to adopt avoidance interpretations of statutes, which gives judges substantial editorial control over statutes, to the point that they sometimes effectively rewrite them. 1998 c. 42 (UK). *See* Ghaidan v. Godin-Mendoza, [2004] 2 AC 557 (HL) 560; *supra* notes 34–39.

202. 474 N.E.2d 567, 570 (N.Y. 1984) (citing N.Y. PENAL L. § 130.35 (1984)).

203. Schachter v. Canada, [1992] 2 S.C.R. 679, 698 (Can.).

text, and the difference between them turns entirely on the particular wording the legislature used.[204]

This insight also extends to the other remedial categories discussed in this Article.[205] There is no meaningful difference, as it concerns minimizing judicial intrusion into the legislative sphere, between as-applied invalidation, facial invalidation, avoidance interpretations, and adding language. Indeed, these categories all blur together in practice. For instance, to strike down an application is effectively to add language to a statute creating an exception.[206] If a court strikes down a statute as applied to group X, that decision is equivalent to adding the clause "this statute shall not apply to group X." The principle distinguishing facial and as-applied challenges is also difficult to discern, since as-applied challenges nearly always define a group larger than just the plaintiff, meaning that a judge must also decide whether the statute applies to parties not before the court in as-applied challenges.[207] And consider avoidance interpretations. Depending on the case, these could also be characterized as adding language to the statute, or as as-applied invalidations. In *Booker*, for example, Justice Stevens' proposed remedy could be understood as interpreting "court" to mean "judge and jury" throughout the Sentencing Reform Act, or as adding the word "jury" to the statute (since Congress clearly intended "court" to mean "judge").[208] The same is true of the opinions in *NAMUDNO*,[209] *Sebelius*,[210] and the many other cases in which the Court effectively rewrote a statute by interpreting it.[211] And the cases that use the avoidance canon to interpret a statute as excluding a constitutionally problematic application could also be described as upholding an as-applied challenge.[212] In short, these remedial categories all run together. If you peer at the distinctions too

204. *See* Duclos & Roach, *supra* note 82, at 34.
205. With the notable exception of severability, which will be treated separately. *See infra* Part V.C.
206. *See* Larry Alexander, *There Is No First Amendment Overbreadth (but There Are Vague First Amendment Doctrines); Prior Restraints Aren't "Prior"; and "As Applied" Challenges Seek Judicial Statutory Amendments*, 27 CONST. COMMENT. 439, 448 (2011); Dorf, *supra* note 110, at 249.
207. *See* Dorf, *supra* note 110, at 294; Fallon, *supra* note 151, at 922–25; Metzger, *supra* note 144, at 882–83.
208. *See supra* notes 202–205.
209. 557 U.S. 193 (2009).
210. 132 S. Ct. 2566 (2012).
211. *See supra* note 171.
212. *See, e.g.*, Bond v. United States, 134 S. Ct. 2077 (2014); Skilling v. United States, 130 S. Ct. 2896 (2010); United States v. X-Citement Video, Inc., 513 U.S. 64 (1994); Int'l Ass'n of Machinists v. Street, 367 U.S. 740 (1961); *see also* Carol Rogerson, *The Judicial Search for Appropriate Remedies Under the Charter: The Examples of Overbreadth and Vagueness*, *in* CHARTER LITIGATION 248 (Robert J. Sharpe ed., 1987) ("While the courts continue to describe reading down as a technique of interpretation rather than of invalidation, as a practical matter reading down is difficult to distinguish from a remedy which would operate to declare particular applications of a law unconstitutional.").

closely, they dissolve into air. How, then, can any of these remedies be considered a greater exercise of legislative power than the others?

To bring this point home, note that judges do not actually change the words of the U.S. Code when they engage in constitutional review—only Congress can do that.[213] All that a court does is change the effects of the law, creating a gap between the law's wording in the statute books and its meaning in the world. Thus to declare certain words invalid, to add words, to strike down applications, or to reinterpret the statute, are all best understood as useful shorthands for judges to communicate how they are changing the statute's meaning in order to fix a constitutional violation. And once we understand that to impose a constitutional remedy is simply to create a gap between textual meaning and real-world effect, then the distinction between these different remedies breaks down entirely. The various remedial categories are simply different strategies for communicating how the statute is to be understood going forward. To treat them as somehow functionally diverse ways of exercising judicial review, or to prefer one to the other because it seems a greater intrusion into the legislative sphere, is to create empty formalist distinctions.

One might push back on this dismissal of Editorial Restraint by defending it on more functional grounds. While each of these remedies may not be inherently any more legislative than the others, it might still be the case that courts exercise more remedial discretion with some types of remedies than with others. For instance, if courts generally had the power to add language to statutes, they would be able to do much more than if they only had the power to strike down language.[214] This observation may help explain our intuition that some of these remedies are greater intrusions into the legislative sphere than others—we look at a judge adding language to a statute, and think "my goodness, that judge can change the statute in so many ways, they have basically replaced the legislature." But the problem with this reasoning is that judges only have the power to change statutes in the narrow context of constitutional review. They cannot generally add language to statutes; they can only do so to solve specific constitutional violations. Thus the scope of remedial power is limited by the problem at hand, not by the nature of the action the judge takes. In *Liberta*, the New York Court of Appeals could not have rewritten the rape statute simply to improve it, for example by eliminating the force requirement, because that would not have

213. *See, e.g.*, 18 U.S.C. § 700 (2012) (establishing criminal penalties for defacing the American flag, notwithstanding *Texas v. Johnson*, 491 U.S. 397 (1989)).

214. *See supra* notes 55–57 and accompanying text.

solved a constitutional violation. Its menu of remedies was limited to the problem before it—that is, making the statute gender neutral.

One might reply that even with this limitation, if judges have the power to add language they will have more options in deciding how to solve the constitutional problem. But there is no reason to think that any one kind of remedy always provides a larger number of options than any other. Depending on the case, a constitutional violation might be solved by striking down one of several provisions, defining one of several possible as-applied exceptions, or adopting one of several possible avoidance interpretations.[215] Choice can exist regardless of the type of remedy. Now admittedly, it is true that in some cases where judges add language to statutes there is no one formulation that flows naturally from the requirements of the Constitution. In such cases the text to be added is indeterminate, and judges must essentially write a complex new provision from scratch.[216] To illustrate, consider the Supreme Court's decision in *Ayotte v. Planned Parenthood of Northern New England*.[217] In that case, the Court held unconstitutional a state law requiring minors to notify their parents before obtaining an abortion. In particular, the Court objected to the lack of an exception for situations where the minor's health is jeopardized. If the Court had decided to add language to the statute carving out such an exception, it would have had to decide questions like what kinds of health risks qualify, how to go about demonstrating the health risks, etc. This would have required some complex legislative drafting by the Court. But this problem is not unique to the context of adding language. If the Court in *Ayotte* opted instead to invalidate the statute as-applied (as it in fact did),[218] the exact same problem would arise. And the same is true if the Court had solved the constitutional violation through an avoidance reading.[219] The Court (or future courts) would still

215. *See, e.g.*, *Skilling*, 130 S. Ct. at 2906 (interpreting a prohibition on "honest-services fraud" as applying only to bribery and kickbacks so as to avoid unconstitutional vagueness—the Court could have added any number of other offenses to this list and still solved the vagueness problem); Ayotte v. Planned Parenthood, 546 U.S. 320, 331 (2006) (finding a parental notification abortion law unconstitutional as applied, and remanding to the lower court to define the class of cases where the law cannot be applied); Bowsher v. Synar, 478 U.S. 714, 783–87 (1986) (Blackmun, J., dissenting) (noting that there are several different provisions that could be struck down to solve the constitutional violation).

216. *See* Schachter v. Canada, [1992] 2 S.C.R. 679, 707 (Can.) ("[T]he court should not read in in cases where there is no manner of extension which flows with sufficient precision from the requirements of the Constitution.").

217. 546 U.S. 320.

218. *Id.* at 332.

219. *Compare* Bond v. United States, 131 S. Ct. 2355 (2011) (interpreting an exception into the Chemical Weapons Convention so that it does not apply to individual acts of poisoning), *with* United States v. Kimber, No. 13-3661-cr (2d Cir. Jan. 30, 2015) (finding that the *Bond* exception does not apply to an individual who targeted a hospital with dangerous chemicals).

have to define the precise scope of the exception. This indeterminacy problem is thus not a reason to generally prefer one type of remedy to another. Instead it is, at least arguably, a reason to opt for a blunter remedy (often striking down the statute) in situations where a rewriting remedy is too complex because the conflict between statute and Constitution could be solved in a large number of ways.[220]

Finally, one might defend Editorial Restraint as formalism for formalism's sake. Even recognizing that there is no deeper reason to think that one type of remedy confers more legislative power on judges than another, one might still embrace a strict hierarchy of remedies simply for the constraint such a hierarchy provides. If judges know that they must first look for an avoidance reading, then an as-applied remedy, then a facial invalidation, and can only add text if none of the above work, they will at least have a remedy-selecting manual that limits their discretion. False scarcity is still scarcity. Of course, any other arbitrary ordering of remedies would achieve the same goal. We could thus mix up the conventional hierarchy, or impose an entirely different one, and still limit judges' discretion. But the conventional hierarchy does at least track our intuitions about which judicial remedies seem more legislative than others, and there may be value, for the sake of public legitimacy, in having judges adhere to those intuitions even if they lack a deeper basis. Editorial Restraint would thus serve as a kind of noble lie, a symbolic performance by judges showing their respect for the limits of their role.[221] The problem with embracing empty formalism, though, is that we readily abandon it in a crisis. It is all well and good for judges to say that they will never add language to a statute when they can strike the statute down instead. But they will not honor that commitment if doing so requires striking down the law against rape, as in *Liberta*. That would be an insane result, and we tend to bend formalisms away from insanity. The Supreme Court even abandoned its rule against adding language in the 1970s sex equality cases, where the alternative was to eliminate public welfare benefits for women.[222] Indeed, the American doctrine of constitutional remedies is full of categorical rules preferring one kind of remedy to another, according to the logic of Editorial Restraint, that are broken when they would cause an undesirable result. The Court announced in *Salerno* that it

220. Though it is worth noting that federal courts deal with indeterminate and complex remedial orders in other contexts, such as cases involving structural injunctions against public institutions like prisons, schools, legislative map drawers, and mental hospitals. *See generally* William A. Fletcher, *The Discretionary Constitution: Institutional Remedies and Judicial Legitimacy*, 91 YALE L.J. 635 (1982).

221. *Cf.* SHAI DOTHAN, REPUTATION AND JUDICIAL TACTICS: A THEORY OF NATIONAL AND INTERNATIONAL COURTS 170–77 (Tom Ginsburg et al. eds., 2015) (suggesting that courts can protect their reputations as neutral decision-makers by pretending that they lack discretion in significant cases).

222. *See supra* notes 120–123 and accompanying text.

categorically prefers as-applied challenges to facial challenges, but this commitment is more honored in the breach than the observance.[223] And the strong rule favoring avoidance readings was broken in *Booker*, where Justice Stevens's preferred avoidance reading would have fundamentally transformed the statute.[224] For those committed to the idea that judges should have as few remedial options as possible in constitutional review cases, Editorial Restraint offers an appealing limit to judges' discretion. The problem, however, is that judges cannot keep themselves tied to the mast. Discretion inevitably reappears when it is wanted.

B. The Case for Purpose Preservation

The basic argument for Purpose Preservation is appealingly simple. The legislature is elected by the voting public to achieve certain goals, and it enacts statutes in service of those goals. A court that interprets and applies these statutes should act as the legislature's (and by extension the voting public's) faithful agent, ensuring that a statute's meaning is not changed in a way that undermines its purposes.[225] A problem, however, arises when the legislature enacts a statute that is partially unconstitutional. The relevant court must then change the statute to fix the constitutional violation. If the court approaches this task as the legislature's faithful agent, it should select the change that best fits with the legislature's goals. Or, framed a slightly different way, it should try to put itself in the collective heads of the enacting legislators and try to imagine which remedy they would hypothetically have preferred. An analogy can be drawn to the interpretation of contracts: just as courts often fill in ambiguous contract terms by determining what the contracting parties would hypothetically have chosen,[226] courts should choose constitutional remedies by determining which fix the enacting legislature would hypothetically have chosen. In so doing, the court will vindicate the legislature's goals, and by extension will also vindicate the goals of the democratic public that voted for the legislature.[227]

223. *See supra* notes 151–153 and accompanying text.
224. *See supra* notes 175–178 and accompanying text.
225. For an analogous discussion of the "faithful agent" concept in the context of statutory interpretation, as a justification for both textualist and purposivist approaches, see Amy Coney Barrett, *Substantive Canons and Faithful Agency*, 90 B.U. L. REV. 109, 112–17 (2010); Manning, *supra* note 49, at 9–21.
226. *See* Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 YALE L.J. 87, 91–92 (1989).
227. Of course, the actual relationship between the policy preferences of voting publics and those of the legislatures they elect is more tenuous than described here. *See* Barry Friedman, *Dialogue and Judicial Review*, 91 MICH. L. REV. 577, 629–43 (1993). But it is not necessary to get into a fully qualified account of representative democracy. The important point is that, because it is elected and must be

There are two related problems with this case for Purpose Preservation. First is the problem of indeterminacy. Statutes are rarely written with instructions for what to do if they are found unconstitutional, and legislative history does not often get into speculation about hypothetical court decisions.[228] Consequently, courts will have to look at the legislative debates over the statute, trying to get a sense of what problem the legislature was responding to, and how it weighed the consequences of the various alternative policies it considered.[229] This can get a bit speculative. In some cases there will be no obviously correct answer at the end of such an inquiry—the court must do the best it can in light of ambivalent evidence. Reasonable minds will differ.[230] And this leads to the second problem: motivated reasoning. If a judge has his or her own views about which remedy is superior as a matter of policy, these can influence how that judge interprets the legislative purpose.[231] The judge might thus search the legislative debate for evidence supporting their position, much like picking one's friends out of a crowd. To illustrate this problem, consider the debate over the remedy in *Booker*. Justices Breyer and Stevens had very different opinions about how Congress would have wanted to fix the Sentencing Reform Act. Justice Breyer thought Congress would have wanted the Sentencing Guidelines made non-mandatory, while Justice Stevens thought it would have wanted to require that each sentencing factor be proven to a jury.[232] Probably not by coincidence, these opinions about Congress's preferences seemed to track each justice's substantive view about which fix was better as a matter of public policy.[233] Now, ambiguity and motivated reasoning will not be problems in every case. In *Liberta* the court could pretty confidently assume that any sane legislature would not want rape legalized. And in equal protection challenges to welfare statutes, where the court can either level up or level down, it is often safe to assume that the legislature would want to level up if the uncovered

reelected, the legislature can legitimately enact statutes to achieve its policy objectives. Purpose Preservation is democratic insofar as it minimizes disruption of such majoritarian policymaking.

228. *See* Fish, *supra* note 23, at 1331; Miller, *supra* note 79, at 89–90.

229. *See* Victoria F. Nourse, *Elementary Statutory Interpretation: Rethinking Legislative Intent and History*, 55 B.C. L. REV. 1613 (2014) (defending the idea of an aggregate congressional "intent").

230. *See, e.g.*, United States v. Booker, 543 U.S. 220, 248 (2005) ("[R]easonable minds can, and do, differ about the outcome . . . .").

231. The same point applies in the context of statutory and constitutional interpretation. *See* Dan M. Kahan, *The Supreme Court 2010 Term—Foreword: Neutral Principles, Motivated Cognition, and Some Problems for Constitutional Law*, 125 HARV. L. REV. 1 (2011).

232. *Booker*, 543 U.S. at 247–49; *id.* at 292 (Stevens, J., dissenting in part).

233. *See* Walsh, *supra* note 8, at 752 ("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."). *Cf.* United States v. Rivera, 994 F.2d 942, 949 (1st Cir. 1993) (Breyer, J.) ("[T]he very theory of the Guidelines system is that when courts, drawing upon experience and informed judgment in such cases, decide to depart, they will explain their departures.").

group is substantially smaller than the covered group, and to level down if the reverse is true.[234] But in ambiguous cases, Purpose Preservation's democratic credentials are undercut. The approach starts to look less like the judiciary channeling the legislative will, and more like the judiciary imposing its own preferences on the statute.

These problems with Purpose Preservation are real. But from the perspective of seeking not to disrupt the legislative will, Purpose Preservation is still better than a formalist approach even in ambiguous cases. This is because a formalist approach, which eliminates judges' ability to discern and apply the legislature's preferences, is far more likely to produce an outcome the legislature would never have considered or wanted. To illustrate, imagine that you have a loyal servant named Lieber. You command Lieber to "fetch some boysenberry ice cream."[235] Lieber arrives at the ice cream store and finds that boysenberry ice cream is not available (because boysenberries have been held unconstitutional). What is Lieber, the faithful agent, to do? If he seeks to maximize your satisfaction, he will pick the ice cream flavor that comes closest to meeting your expressed preferences. Since you sought a berry flavored ice cream, perhaps raspberry, blueberry, or strawberry would be your next choices. Certainly pistachio or rum raisin will not do. And within the universe of berry flavored ice creams, Lieber's own preferences might sneak into his reasoning. If he happens to think that raspberry is the most delicious berry flavor then he may, in the face of ambiguity, opt to purchase that flavor. And this is perfectly consistent with Lieber's self-understanding as a loyal agent. He could reason "raspberry is the best of the berry flavors, so I believe it would have been my employer's second choice." You may prefer blueberry to raspberry, but unless you tell Lieber as much, he will have to infer your preferences from limited evidence, and so his own preferences may influence his selection.

Now imagine that you fire Lieber for bringing back raspberry ice cream, and replace him with a robot: the Ice Cream Fetchbot 5000. Fetchbot is programmed to go to the store and fetch the flavor of ice cream you ask for. But unlike Lieber, Fetchbot lacks the human ability to get into your head and reason through what would have been your second choice. So Fetchbot must be programmed with certain fallback rules if the flavor you ask for is unavailable. It might be programmed to grab the next flavor in alphabetical order, to grab the least expensive flavor, to grab a flavor at random, or to come back with no ice

234. *See* Schachter v. Canada, [1992] 2 S.C.R. 679, 712 (Can.).

235. *Cf.* FRANCIS LIEBER, LEGAL AND POLITICAL HERMENEUTICS, OR PRINCIPLES OF INTERPRETATION AND CONSTRUCTION IN LAW AND POLITICS 18 (1837), *as reprinted in* 16 CARDOZO L. REV. 1883, 1904–05 (1995) (offering a similar example).

cream at all. Like a good formalist judge, Fetchbot will follow these instructions without exercising any creative reasoning. Is Fetchbot really better than Lieber at satisfying your preferences? Formalist approaches, like Editorial Restraint, make judges into Fetchbots. They give judges an unbending set of rules for selecting remedies: strike down language rather than adding it, strike down an application rather than an entire statute, adopt an avoidance reading rather than striking anything down. But if these rules produce the result the legislature would have wanted, it is only by chance. Having a judge consciously try to discern the hypothetical legislative preference, despite the risks of indeterminacy and motivated reasoning, is more likely to produce a result consistent with the legislature's goals.

The other alternative to Purpose Preservation would be the Norm Recognition approach discussed in Part II.C. Under such an approach, judges would scour the entire legal-political landscape—the Constitution, statutes, judicial doctrines, agency rules, etc.—for norms to apply, and would select among these norms the one that seems the most powerful in the particular context of the case. Ronald Dworkin articulates this approach in the context of legal interpretation, but it applies just as well in the context of constitutional remedy selection.[236] Whatever else one might say about it, Norm Recognition does not preserve majoritarian democracy. Instead, the value it serves is coherence—judges constructively interpret the legal system as embodying a discernable hierarchy of values, and so impose coherence where none might otherwise exist.[237] Norm Recognition is in fact peculiarly undemocratic, since it empowers judges to decide deep value questions about our system. And the problems of indeterminacy and motivated reasoning are worse for Norm Recognition than for Purpose Preservation: finding the legislature's preferences is a more tractable inquiry than finding the implications of competing abstract norms. It is difficult to imagine that a judge's own substantive moral and legal views would not influence their selection of norms. Norms are wonderfully malleable things, and can often be read to justify any result or its opposite.[238]

Purpose Preservation has its problems. But when considered against the main alternatives, it emerges as the best option. Formalist approaches like Editorial Restraint successfully constrain judges' discretion, but they do so at the cost of satisfying legislative preferences (and, by extension, the preferences of the democratic public). And Norm Recognition replaces the legislature's preferences with

236. *See* DWORKIN, *supra* note 107; *see also* Sunstein, *supra* note 74, at 436.

237. *See* DWORKIN, *supra* note 107, at 245; *see also* EINER ELHAUGE, STATUTORY DEFAULT RULES: HOW TO INTERPRET UNCLEAR LEGISLATION 31–35 (2008).

238. *See* Bizar, *supra* note 52, at 158; *see also* K.N. Llewellyn, *The Normative, the Legal, and the Law-Jobs: The Problem of Juristic Method*, 49 YALE L.J. 1355, 1359–60 (1940).

judges' own views about competing abstract norms, rendering the problems of indeterminacy and motivated reasoning far worse. Judges should understand themselves as intelligent servants of the democratic will when selecting constitutional remedies. By doing so, they will best balance the need to vindicate popular preferences against the danger of replacing those preferences with their own views.

V. REFINING PURPOSE PRESERVATION

Now that Purpose Preservation has been defended as the best approach to choosing constitutional remedies, it remains to show how it might work in practice. This final Part explores the logic of Purpose Preservation in a bit more detail, considering how judges should use it and noting several important caveats.

A. Whose Purpose?

Exactly whose purposes are being preserved? Since the argument for Purpose Preservation is that it fits better with majoritarian democracy, the ultimate answer must be those of the voting public. Courts, however, cannot practically discover the public's preferences concerning a specific constitutional remedy.[239] Courts must therefore look instead to the public's representatives—the elected legislators. But should courts consider the views of the current legislature, or those of the legislature that enacted the challenged law? Einer Elhauge has argued, in the statutory interpretation context, that the preferences of the current legislature should control the meaning of ambiguous statutes.[240] This approach makes a certain kind of sense—the current legislature, after all, is more likely to reflect the current popular will than is a past legislature.[241] And by the same logic,

239. Even aside from the practical difficulties of conducting an opinion poll, the public would presumably not have coherent views on the particular case at bar.

240. ELHAUGE, *supra* note 237, at 40–46; *see also* Beers, *supra* note 125, at 120–21.

241. Elhauge also makes a quasi-Rawlsian argument for using the current legislature's preferences to decide the meaning of ambiguous statutes. He claims that any legislature, when faced with a choice between the "current legislature" rule and the "enacting legislature" rule, would opt for the former because that rule would maximize its influence. ELHAUGE, *supra* note 237, at 43. The legislature would get to control the meaning of all the laws on the statute books for two years (or whatever the interval is between elections) rather than controlling the meaning of the small number of statutes that it enacts for all time. *Id.* at 43–46; *see also* Einer Elhauge, *Preference-Estimating Statutory Default Rules*, 102 COLUM. L. REV. 2027, 2086–95 (2002). I am not sure I agree with Elhauge's reasoning. If I were a legislator faced with that choice, my decision would probably depend on whether the legislature I served in had enacted any major laws. If it had not, I would opt for two years of controlling the meaning of every law, but if it had, I might instead opt for controlling the meaning of the laws we enacted for all future time. And if I were a member of the Democratic Party, I would

one could argue that the preferences of the current legislature should control how an unconstitutional statute is fixed. However, there are two reasons that judges should instead look to the views of the enacting legislature. First, the enacting legislature is far more likely to have coherent preferences with respect to the statute in question. Second, substituting the views of the current legislature undermines the principle of intertemporal democracy—it harms a past electoral majority by refusing to preserve its legislative accomplishments, and grants an unfair windfall to the current electoral majority by imposing its preferences on the product of a past majority.

The first problem with the "current legislature" approach is that it is unlikely that the current legislature has coherent or discernable views about the statute in question. With the enacting legislature, we know that it must have had specific ideas about the statute because it passed the thing. Thus even if there is no legislative history whatsoever, we can still look at the terms of the statute and the problem it was designed to solve, and infer from these the goals the enacting legislature had in mind. The current legislature, however, may have no discernible aggregate views about a statute enacted in the past. A legislature is made up of many different individuals, all with their own opinions, and if it does not speak as a unit by enacting something there is no coherent way to ascribe to it views about a particular law.[242] This is especially true in a system like ours, where different political parties often simultaneously control different branches of government.[243] Consider the constitutional challenge to the Affordable Care Act (ACA). The ACA was enacted in 2010 on a party-line vote—the Democrats controlled the House, had a filibuster-proof majority in the Senate, and had Barack Obama in the White House.[244] But by the time a constitutional challenge to the ACA reached the Supreme Court in 2012, the House had switched to Republican control and voted several times to repeal the

probably opt for the "enacting legislature" rule, given that most of the major statutes enacted in the last century (e.g. the Civil Rights Act, Voting Rights Act, National Labor Relations Act, Administrative Procedure Act, Affordable Care Act, etc.) were passed by congresses controlled by the Democrats. But the question is not which rule would be chosen by a legislature looking to maximize satisfaction of its own preferences vis-à-vis past and future legislatures. Instead, the question is which rule would best fit within the logic of America's system of representative democracy.

242. *Cf.* Shepsle, *supra* note 74, at 249 (arguing that even when a legislature does act collectively, it lacks a single collective will).

243. The same party has only simultaneously held the House, Senate, and Presidency for roughly one-third of the years since 1951.

244. Shailagh Murray & Lori Montgomery, *House Passes Health-Care Reform Bill Without Republican Votes*, WASH. POST (Mar. 22, 2010), http://www.washingtonpost.com/wp-dyn/content/article/2010/03/21/AR2010032100943.html [http://perma.cc/YB5J-BCLA].

ACA.[245] Meanwhile, the Democrats still controlled the Senate (albeit without a filibuster-proof majority) as well as the presidency. How should the Supreme Court have applied the "current legislature" approach in these circumstances? One chamber wanted to repeal the law *in toto*, a bare majority of the other still supported it (but was constrained by the filibuster rule), and the President strongly supported it. The Supreme Court could not meaningfully look at these three government bodies in 2012 and decide whether to strike down the entire ACA or just part of it based on current legislative preferences. There was no single coherent set of current preferences. And even if the issue had not been so clearly polarized, there would still be no way of discerning what the current government wanted. We can only see congressional preferences when Congress actually acts.[246]

The second problem with the "current legislature" approach is that it undermines what could be called intertemporal democracy—the idea that different majority coalitions exercise power at different times. In our political system, one majority can be swept into power in one election only to be replaced by a different majority a few years later. But the fact that a party loses power does not mean that its agenda has been totally repudiated by the voting public. Rather, shifts in political power cause different coalitions of voters to take turns having their views reflected in the legislative majority.[247] In this system, each new government must use its scarce years in power to pursue a limited legislative agenda. If agenda-setting resources were infinite, then every new majority could rewrite all the laws according to its preferences. But each election is not the French Revolution—we do not burn all the deeds and reset the calendar to zero. When coalition A is in the majority, it makes at most a handful of significant changes to the legislative code, and when coalition B is in power a few years later, it does the same.[248] Looking at democracy across time in this way, one sees that it is only fair to have coalition A's preferences control the laws it passed, and to do

245. *See* Ed O'Keefe, *The House Has Voted 54 Times in Four Years on Obamacare. Here's the Full List.*, WASH. POST (Mar. 21, 2014), http://www.washingtonpost.com/blogs/the-fix/wp/2014/03/21/the-house-has-voted-54-times-in-four-years-on-obamacare-heres-the-full-list/ [https://perma.cc/6QL9-VNCH].

246. *See* Bizar, *supra* note 52, at 146 ("The only way for the court to really know how the current legislature would amend the statute is to wait and see the amendment.").

247. This idea is closely connected to Robert Dahl's idea of "polyarchy," a system where majorities are made up of constantly shifting coalitions between different minority groups. *See generally* ROBERT A. DAHL, A PREFACE TO DEMOCRATIC THEORY (1956); ROBERT A. DAHL, POLYARCHY (1971).

248. *See* DAVID R. MAYHEW, DIVIDED WE GOVERN: PARTY CONTROL, LAWMAKING, AND INVESTIGATIONS, 1946–2002, at 206–13 (2d ed. 2005) (showing that from 1991 to 2002, each two-year Congress averaged roughly ten significant enactments).

the same for coalition B.[249] Coalition A expended limited resources to enact the laws it passed, and the subsequent election of coalition B does not in itself mean the public repudiated coalition A's achievements. Further, it would give coalition B an unfair windfall to impose its remedial preferences on the challenged law, since coalition B has not expended the time and resources needed to amend or repeal that law. Consider once more the ACA. The fight to pass the ACA dominated the first two years of Barack Obama's presidency, and the Democratic majority elected in 2008 accomplished little else despite controlling both chambers of Congress.[250] If a court considering a constitutional challenge to the ACA looked to the preferences of a later, Republican-controlled Congress, and used those preferences to strike down the ACA in its entirety, this would be unfair to the 2008 Congress and the voting coalition that elected it.[251] It would also be an unjust windfall to the later Congress—if that Congress cannot repeal the statute legislatively, the courts should not do the job for it through constitutional remedies.[252]

There is, however, an important exception to this rule favoring the preferences of the enacting legislature. If the challenged law was passed a long time ago, and is now anachronistic (but persists due to legislative inertia), then courts should not focus on the will of the enacting legislature. The case for Purpose Preservation relies on the argument that it enhances majoritarian democracy, and this argument breaks down if the challenged law's purposes are so dated that they have lost any claim to popular support.[253] Such loss of support cannot happen in

249. An interesting question arises in cases where a court can strike down either one of two provisions, and each provision was enacted by a different Congress. *See, e.g.*, Bowsher v. Synar, 478 U.S. 714, 783–87 (1986) (Blackmun, J., dissenting) (considering whether to strike down a provision enacted in 1985 granting the Comptroller General certain powers, or a provision enacted in 1921 providing for removal of the Comptroller General by joint resolution); *see also* Free Enterprise Fund v. Pub. Co. Acct'g Oversight Bd., 561 U.S. 477 (2010) (holding that having two layers of for-cause removal protection for an accounting oversight board, one for officers of the Securities and Exchange Commission and one for members of the board itself, violates the President's removal power). In such cases, the court must somehow compare the two different legislatures' hypothetical preferences. One way to do this might be to imagine that the legislatures were brought together with the aid of a time machine, and to reason through what remedy this intertemporal joint session of two congresses would settle on.

250. *See generally* STAFF OF THE WASH. POST, LANDMARK: THE INSIDE STORY OF AMERICA'S NEW HEALTH-CARE LAW AND WHAT IT MEANS FOR US ALL 11–62 (2010).

251. *See* Fish, *supra* note 23, at 1351.

252. The "current legislature" approach would also have the strange feature that the selection of remedies would change radically depending on which party is in power. This is somewhat unseemly. Among other things, it would cause impact litigation groups to strategically wait for a friendly Congress before bringing constitutional challenges.

253. A more difficult question is what to do in cases where a statute's original purposes are anachronistic, but only because the statute's meaning has changed over time through judicial interpretation. For example, the Sherman Act was passed in large part to protect small businesses from larger ones, but

only a few elections, but can happen over a longer stretch of time if public values have clearly shifted away from those embodied in the challenged law. Several scholars have made similar arguments in other contexts. William Eskridge has argued that where a statute's meaning is ambiguous and the original legislative purpose anachronistic, the statute should be "dynamically" interpreted in light of current values.[254] And Guido Calabresi has argued that judges should have the power to strike down anachronistic laws, but only provisionally, leaving the legislature able to reenact them.[255] By similar reasoning, courts committed to Purpose Preservation might substitute what they take to be current public values in cases where the enacting legislature's goals are clearly out of place in time. For example, Texas law used to include the following provision: "Homicide is justifiable when committed by the husband upon one taken in the act of adultery with the wife, provided the killing takes place before the parties to the act have separated."[256] Imagine a case in which a woman is accused of murdering her spouse's lover and argues that the exclusion of wives from this provision violates the Equal Protection Clause, and further that it should be leveled up to apply to her.[257] What should the court do? If it looked to the views of the law's enactors to solve this problem, it might well conclude that the relevant Texas legislature would have preferred to let wives kill adulteresses rather than imprisoning men for killing adulterers. But the statute's purpose reflects a highly dated concept of marriage, in which the wife is the husband's property and he can commit violence to enforce his possession of her.[258] Recognizing this anachronism, a court might choose instead to level the statute down so that neither men nor women could invoke it as a defense to murder. By so doing the court would not be undermining

over the last half-century legal elites in the academy and on the bench have effectively changed its purpose to be about maximizing consumer welfare. *See* Christopher Grandy, *Original Intent and the Sherman Antitrust Act: A Re-examination of the Consumer-Welfare Hypothesis*, 53 J. ECON. HIST. 359 (1993). If a constitutional challenge were brought against the Sherman Act, should the court choose a remedy based on its original purpose or this changed purpose? Tentatively, one way to resolve this problem might be through a sort of adverse possession theory. The legislature has had several decades to reverse this shift in the Act's meaning, and has not done so. One might see this as a tacit majoritarian acquiescence to the change.

254. WILLIAM N. ESKRIDGE, JR., DYNAMIC STATUTORY INTERPRETATION 10–11 (1994).

255. *See* GUIDO CALABRESI, A COMMON LAW FOR THE AGE OF STATUTES (1982).

256. TEX. CRIM. STAT. § 1220 (1925). This provision remained in force until nearly 1970. *See* JOEL FEINBERG, DOING AND DESERVING 102 (1970).

257. This argument was in fact made by a defendant named Pearl Reed in 1933. *See* Reed v. State, 59 S.W.2d 122, 123 (Tex. Crim. App. 1933). The Texas appeals court upheld Reed's conviction. *Id.* at 124.

258. *See* CYNTHIA LEE, MURDER AND THE REASONABLE MAN: PASSION AND FEAR IN THE CRIMINAL COURTROOM 21 (2003).

the principle of intertemporal democracy, since such a law lacks a claim to popular support in modern society.

B. Inter-systemic Review, Irony Avoidance, and Inherent Legislative Powers

Purpose Preservation need not be adopted in pure form. It can also be modified to account for principles beyond the need to vindicate the legislature's goals. The following are three potential modifications to Purpose Preservation that would allow it to incorporate important extrinsic principles.

First, the argument for Editorial Restraint has more force in cases where the courts of one government review the laws of another. This is because, as shown in Part I, the selection of constitutional remedies is a quasi-legislative decision: a court actually changes the statute's meaning in order to make it consistent with higher law. And the power to define one's own laws is a fundamental aspect of state sovereignty. Thus, for example, the European Court of Human Rights and the Inter-American Court of Human Rights cannot change a country's law even if they find that the law violates a human rights treaty—they are restricted to awarding damages to the party whose rights were infringed.[259] Similarly, in the United States, federal courts that review the constitutionality of state statutes often decline to impose remedies themselves, deferring instead to state courts.[260] For example, in *Skinner v. Oklahoma* a man was convicted of larceny, and the state of Oklahoma sentenced him to be sterilized.[261] He brought a constitutional challenge in the Supreme Court, arguing that the law providing for sterilization of larcenists was unconstitutionally irrational because it did not do the same for embezzlers. The Court agreed, but remanded the case to the Oklahoma Supreme Court to determine the proper remedy—whether to expand the statute to also sterilize embezzlers, or to stop sterilizing larcenists (blessedly, it chose the latter).[262] The Court has often similarly remanded to state courts cases involving severability questions or the choice between leveling a law up or down.[263] Sending the remedial question back to the state court is an ideal solution in such cases. If the Supreme Court instead attempted its strategy from *Northern Pipeline*—using

259. *See* Barbra Fontana, *Damage Awards for Human Rights Violations in the European and Inter-American Courts of Human Rights*, 31 SANTA CLARA L. REV. 1127, 1127 (1991).

260. *See* Abbe R. Gluck, *Intersystemic Statutory Interpretation: Methodology as "Law" and the* Erie *Doctrine*, 120 YALE L.J. 1898, 1948–58 (2011).

261. 316 U.S. 535, 536 (1942).

262. Skinner v. State, 155 P.2d 715, 716 (Okla. 1945).

263. *See, e.g.*, Hooper v. Bernalillo Cty. Assessor, 472 U.S. 612 (1985); Zobel v. Williams, 457 U.S. 55 (1982); Wengler v. Druggists Mut. Ins. Co., 446 U.S. 142 (1980); Orr v. Orr, 440 U.S. 268 (1979); Stanton v. Stanton, 421 U.S. 7 (1975).

delayed invalidation to induce the relevant legislature to craft its own solution—there would be no way of guaranteeing that the state legislature would act.[264] A state court, by contrast, can be expected to actually choose a remedy. However, a problem arises in cases where litigation over a state statute is brought first in federal court. In such cases, remand to a state court is not possible. This problem has led the Supreme Court to choose remedies on its own in cases where it would have better served federalism principles to remand to a state supreme court.[265] In such cases, the Supreme Court (or even a lower federal court) should take advantage of the certified question procedure, whereby it can ask a state supreme court to resolve a discrete question of state law.[266] Use of the certification procedure would let the Supreme Court find a state law unconstitutional without having to infringe on the state's sovereignty by choosing how the law is changed.

A second kind of case where the argument for relying on the legislative purpose weakens is one where that purpose conflicts with the purpose of the constitutional provision being enforced. This is the principle of Irony Avoidance discussed in Part II—that courts should not choose remedies that undermine the right being protected. Irony Avoidance can be incorporated into Purpose Preservation as an exception to the general rule that the goals of the enacting legislature control the choice of remedy. Such an exception actually fits with the logic of Purpose Preservation—it just expands the relevant purposes to include those of the trumping constitutional provision, instead of only the statute. Indeed, the Canadian Supreme Court in *Schachter* embraced such an exception while articulating its overarching commitment to Purpose Preservation. The opinion declared that, when deciding whether to expand or strike down an underinclusive law, "respect for the role of the legislature and the purposes of the *Charter* are the twin guiding principles."[267] This constitutional irony exception helps us make sense of situations where the legislative purpose is itself unconstitutional.[268]

264. *See supra* notes 185–92 and accompanying text.

265. *See, e.g.*, Dorchy v. Kansas, 264 U.S. 286, 291 (1924); Ginsburg, *supra* note 22, at 313. *Compare Stanton*, 421 U.S. at 18 (a case arising from state court, where Supreme Court remanded for remedy), *with* Craig v. Boren, 429 U.S. 190, 210 n. 24 (1976) (a similar case arising from lower federal court, where Supreme Court imposed a remedy itself and noted that the state legislature could select a different remedy).

266. *See* Guido Calabresi, *Federal and State Courts: Restoring a Workable Balance*, 78 N.Y.U. L. REV. 1293, 1306 (2003).

267. Schachter v. Canada, [1992] 2 S.C.R. 679, 715 (Can.); *see also id.* at 701 (citing Caminker, *supra* note 11).

268. Of course, any statute that violates the Constitution at some level has a purpose that is unconstitutional. But a relatively straightforward distinction can be made between situations where the primary purpose of the statute is unconstitutional (e.g. enacting a law to preserve racial segregation) and situations where the constitutional violation is in service of otherwise valid goals

Take for example a case in which a legislature enacts a law intended to establish racial segregation.[269] If a court finds part of that law unconstitutional, it makes little sense to choose the remedy that best promotes the desire to preserve segregation. This is true even if there is a remedy that would render the statute formally constitutional despite its evil intent.[270] Since the constitutional provision formally trumps the statute as a matter of law, it makes sense that the constitutional purpose should trump the statutory purpose as a matter of remedy selection.

Finally, some judges and commentators have argued that courts should not do certain things through remedial orders, such as expanding criminal liability,[271] or forcing the legislature to budget more money,[272] for fear of infringing on the legislature's fundamental prerogatives. According to this argument, Purpose Preservation has certain limits—some remedies should not be imposed even if they best fit with the legislature's goals.[273] This is not a reason for preferring one particular kind of remedy over another as a general matter. Criminal liability or budgetary obligations could be expanded through any of the standard remedies: adding language, striking language down, striking down applications, or adopting an avoidance reading.[274] Rather, it is a reason to avoid remedies that cause certain specific results, like criminalizing more conduct or spending more money. The problem with this argument, however, is that these issues are not unidirectional. If criminal law and budgetary matters are uniquely within the province of the legislature, then both expanding and contracting criminal liability or government spending will infringe on the legislative domain. And in cases

(e.g. regulating electoral speech in order to combat corruption in politics). *See* Dorf, *supra* note 110, at 279.

269. *See, e.g.*, Griffin v. Cty. Sch. Bd. of Prince Edward Cty., 377 U.S. 218, 221 (1964).

270. *See Schachter*, [1992] 2 S.C.R. at 703 ("[W]here the purpose of the legislation is itself unconstitutional, the legislation should be struck down in its entirety. Indeed, it is difficult to imagine anything less being appropriate where the purpose of the legislation is deemed unconstitutional . . . .").

271. *See, e.g.*, Beers, *supra* note 125, at 131–32.

272. *See, e.g.*, Califano v. Westcott, 443 U.S. 76, 95 (1979) (Powell, J., concurring in part and dissenting in part); *Schachter*, [1992] 2 S.C.R. at 709–10; Beers, *supra* note 125, at 136–40.

273. The court in *Schachter* also suggested that a court should not impose a remedy if the legislature specifically rejected the policy that that remedy creates. *See* [1992] 2 S.C.R. at 708–09. I disagree with this suggestion. If a legislature is debating between two versions of a statute, X and Y, and ultimately settles on X, Y might still be its fallback preference if X is unconstitutional. Indeed, the very fact that it debated Y means that it considered the possibility. In *Booker*, for example, Justice Breyer's remedy (making the Guidelines advisory) was considered and rejected by the enacting legislature. *See* United States v. Booker, 543 U.S. 220, 292 (2005) (Stevens, J., dissenting in part). But notwithstanding that fact, it could still have been the remedy that was closest to what Congress wanted.

274. *See Schachter*, [1992] 2 S.C.R. at 709.

where the existing law is unconstitutional, the court must generally make a change that will either expand or contract the law. In *Liberta*, for example, expanding the rape statute to solve the equal protection violation is not more of an infringement on legislative prerogatives than striking the law down.[275] Similarly, in cases with underinclusive entitlement statutes, forcing the legislature to spend more money is no worse than preventing it from spending money that it has budgeted.[276] These are surely core areas of legislative power, but when the legislature acts unconstitutionally the courts are forced to impose a remedy. They cannot avoid the problem that their decisions will change the existing policy—the best they can do is try to approximate legislative preferences.

C. The Special Case of Severability

As shown in Part I, judges have discretion in choosing constitutional remedies simply because multiple options exist. A judge must find an adequate remedy when a statute violates the Constitution, and if there is more than one change that will fix the violation, then (absent legislative instruction) the judge must choose from amongst the possible changes. Judges do not seek out this discretion—it is thrust upon them. But the same reasoning does not apply in severability cases, because in those cases the statute has already been made constitutional through some other remedy. Severability determinations involve the further question of what should then happen to the rest of the statute. There is no defensible basis for a broad remedial power to declare laws inseverable.[277] Courts do not have a general power to invalidate constitutionally valid statutory provisions, and there is no reason why finding part of a statute unconstitutional would create such a power. None of the three theories that could justify a Purpose Preservation-based severability doctrine—(i) that judges have a delegated remedial power to alter partially unconstitutional statutes, (ii) that after striking down part of a statute, judges can "interpret" the rest in a way that removes further provisions, or (iii) that judges enforce legislative deals in a way analogous to contracts—is defensible in the federal system.[278]

We might nonetheless imagine a system where severability decisions are discretionary—Congress could potentially delegate to courts a power to make

275. *See supra* notes 1–5 and accompanying text. Judicial expansion of criminal laws may conflict with defendants' constitutional rights, but that is a separate issue. *See supra* note 92 and accompanying text.

276. *See supra* notes 120–123 and accompanying text.

277. *See* Fish, *supra* note 23. However, the remedial theory is more plausible in some state systems. *Id.* at 1352–55.

278. *Id.* at 1319–32.

legislation inseverable. In such a system, the Purpose Preservation approach would be substantially more troubling than it is for the other kinds of remedial questions discussed in this Article.[279] This is because judges in severability cases are not limited to solving a discrete constitutional violation. Instead, they are tasked with a broader mandate to alter the law to fit with the enacting legislature's preferences. After a court struck down part of a statute, it could then search through the rest of the statute (or perhaps even the rest of the code) for other things to strike down according to its understanding of the legislature's goals.[280] In *Sebelius*, for example, if the Supreme Court had struck down the individual mandate, it could have gone through the dozens of other provisions in the Affordable Care Act and decided what should stay and what should go based on its estimation of Congress's preferences.[281] Courts could also, arguably, even add language to partially unconstitutional statutes to further the legislative purpose. While it is generally assumed that courts can only invalidate language in severability cases, there is no principled reason for this restriction, since courts can add language to remedy constitutional violations.[282] Indeed, in *Booker*, Justice Breyer effectively added language to the Sentencing Reform Act by providing for a new kind of appellate review—"reasonableness" review—that would fit better with the post-*Booker* statute.[283] Because judges in severability cases are not limited by the task of solving a specific constitutional problem, they behave like legislators when reformulating a law to preserve Congress's goals. They do not choose from a limited set of options to solve a single constitutional problem, but rewrite the whole statute to conform it to Congress's hypothetical preferences.

For these reasons, something akin to Editorial Restraint is preferable in severability cases. Judges should not be empowered to preserve the purpose of a partly unconstitutional statute through further remedial alterations. Instead, they should be restricted to only finding further provisions inseverable in circumstances where the statute itself makes such a finding clear. Severability should thus be treated as a

279. A number of academics argue that severability of statutory language and severability of applications present the same issues. *See supra* note 144. But this is not quite right. One important difference is that, while facial and as-applied challenges are two different ways of resolving constitutional violations (one by striking down an application, the other by striking down an entire provision), a finding of inseverability takes place after the constitutional violation has already been fixed. A second difference is that, while multiple applications can only be grouped through a facial challenge if they are all covered in the same textual provision, an inseverability finding can be applied to any other textual provision elsewhere in the statute.

280. *See* Fish, *supra* note 23, at 1313–19.

281. Nat'l Fed'n of Indep. Bus. v. Sebelius, 132 S. Ct. 2566 (2012).

282. *See supra* Part III.A; *see also* Smith, *supra* note 140 (arguing for an approach to severability in legislative veto cases that would entail adding language).

283. 543 U.S. 220, 261–62 (2005); *see also* Metzger, *supra* note 144, at 890.

matter of statutory interpretation, not a matter of remedy. I have defended one such approach—"severability as conditionality"—which would instruct judges to only find provisions inseverable if there is an explicit inseverability clause, or if those provisions are made textually incoherent or impossible to enforce once the unconstitutional provision is struck down.[284] Several other commentators have defended similar approaches.[285] These restrained, interpretive approaches to severability are actually more consistent with modern severability cases than is the Supreme Court's official Purpose Preservation-based doctrine.[286]

D. The Problem With Provoking Overrides

When a court chooses a constitutional remedy, its choice is always provisional—the legislature can impose a different remedy by amending the statute. For example a court might level down an underinclusive statute, only to have the legislature respond by leveling the statute up. Such legislative solutions are preferable, in a way, because they eliminate the problem of having the court act as a surrogate legislature. For this reason, judges sometimes actually solicit the legislature to choose its own remedy. They do so in a number of ways. Sometimes judges call on the legislature to fix the problem, as the English courts do when they make declarations of incompatibility,[287] and as Justice Breyer did in *Booker* when he wrote "[o]urs, of course, is not the last word: The ball now lies in Congress' court."[288] Judges also sometimes delay their chosen remedy to give the legislature time to substitute its own, as the Canadian Supreme Court did in *Schachter*,[289] and as the United States Supreme Court tried to do in *Northern Pipeline*.[290] These methods are relatively innocuous—they involve judges prodding the legislature, not forcing the legislature's hand.[291]

284. *See* Fish, *supra* note 23, at 1343.
285. *See, e.g.*, Dorsey, *supra* note 137; Rachel J. Ezzell, Note, *Statutory Interdependence in Severability Analysis*, 111 MICH. L. REV. 1481 (2013); Walsh, *supra* note 8. Tom Campbell argues for the opposite approach—that everything should be made inseverable without exception. *See* Campbell, *supra* note 54.
286. *See* Fish, *supra* note 23, at 1358; *supra* notes 139–141 and accompanying text. However the dissent in *Sebelius* suggests that at least some Supreme Court justices might favor a more aggressive approach. *See* 132 S. Ct. at 2642–77 (Scalia, Kennedy, Thomas & Alito, JJ., dissenting).
287. *See supra* notes 40–46.
288. 543 U.S. at 265.
289. Schachter v. Canada, [1992] 2 S.C.R. 679, 725 (Can.).
290. 458 U.S. 50, 87–88 (1982).
291. Though it is worth noting that, in England and Canada, these "prods" are seen as somewhat more coercive. *See supra* note 49 and accompanying text; *Schachter*, [1992] 2 S.C.R at 716–17 (discussing how delaying a remedy forces the issue onto the legislative agenda).

Another strategy, however, is rather more heavy handed. In the statutory interpretation context, some scholars have argued that judges should interpret a statute so as to deliberately subvert the legislature's preferences, thus forcing the legislature to act.[292] Einer Elhauge labels this approach a "preference-eliciting statutory default rule"—rather than mimicking legislative preferences, the court tries to force the legislature to reveal them (either ex ante in the text of the statute or ex post through an override).[293] Such an approach could also work in the context of choosing constitutional remedies. Rather than selecting the best remedy from the perspective of the enacting legislature, a court could instead select the worst remedy from the perspective of the current legislature. The court would thereby effectively force the current legislature to override this remedy and impose its own. There are two major problems with this strategy. First, in the American system it would predictably lead to situations where the court imposes a disastrous remedy and the legislature is unable to muster an override. Second, it conflicts with the principle of intertemporal democracy by undermining both the enacting and the current legislature's power to define its own agenda.

If a court is going to deliberately choose a remedy that the legislature will dislike, it should be confident that the legislature will actually override this choice. But there is no reason to think that judges are good at predicting legislative action. Recall the *Northern Pipeline* saga. The Supreme Court delayed its remedial order twice anticipating a legislative response, and after it became clear that no response was forthcoming the judiciary itself had to step in with a solution.[294] The stakes in that case were not exactly low—the Court's remedial order destroyed the existing system of bankruptcy adjudication, and Congress still failed to act. Such legislative inertia is a basic feature of our system, with its numerous vetogates and high degree of partisanship.[295] One simply cannot be confident that Congress will avert a judicially created policy crisis.[296] And the

292. Ayres and Gertner first developed this concept in the context of contract interpretation. *See* Ayres & Gertner, *supra* note 226. Elhauge and Eskridge then expanded it to statutory interpretation. *See* ELHAUGE, *supra* note 237, at 151–67; Elhauge, *supra* note 241; ESKRIDGE, *supra* note 254, at 151–61 (arguing for an interpretive canon in which the group with the least access to political power wins in close cases).

293. *See* Einer Elhauge, *Preference-Eliciting Statutory Default Rules*, 102 COLUM. L. REV. 2162 (2002).

294. *See supra* notes 185–195 and accompanying text.

295. *See* William N. Eskridge, Jr., *Vetogates,* Chevron*, Preemption*, 83 NOTRE DAME L. REV. 1441 (2008); Richard H. Pildes, *Romanticizing Democracy, Political Fragmentation, and the Decline of American Government*, 124 YALE L.J. 804 (2014).

296. *See, e.g.*, CALABRESI, *supra* note 255, at 269–70 (discussing Congress's failure to provide for criminal jurisdiction over military dependents abroad, after the Supreme Court struck down the statute providing such jurisdiction in *Reid v. Covert*, 354 U.S. 1 (1957)—this left such dependents free to commit crimes with impunity); Jamelle Bouie, *The GOP's Sick Maneuvers: If the Supreme Court Throws Obamacare Into Chaos, Republicans Are Content to Just Watch It Fail*, SLATE (Feb. 3, 2015,

problem of inertia has gotten significantly worse over the last few decades. As William Eskridge and Richard Hasen have shown, legislative overrides of the Supreme Court's statutory interpretation decisions have fallen off dramatically since 1998.[297] Given this reality, choosing a remedy to provoke an override is a risky strategy.[298] We could wind up with a law that neither the enacting nor the current legislature would have wanted, and that the court chose specifically because it seemed like a bad idea.

Even if a court could be certain of provoking an override, preference-eliciting remedies would still be problematic because they undermine the principle of intertemporal democracy—that each legislative majority should be able to define the policy agenda for the period it governs. In our system different majority coalitions will be elected at different times, and each will have its own set of legislative priorities that matter to the people who elected it. For a court to force an override infringes on the agenda-setting power of both the enacting and the current legislature. Imagine, for example, that Congress I cares a great deal about endangered animal species, and so devotes a lot of time and resources to enacting the "Endangered Species Act." A few years later Congress II is elected, and Congress II is weakly opposed to protecting endangered species, but does not consider repealing the Endangered Species Act a priority. While Congress II is in power, a constitutional challenge is successfully brought against the Endangered Species Act, and the court must choose among several constitutional remedies. It picks a remedy designed to provoke an override from Congress II, and Congress II responds by repealing the entire law. This tactic harms the democratic process with respect to both congresses. Congress I (and the voters that elected it) considered protecting endangered species a major priority, and devoted a lot of its agenda-setting resources to passing the law. The court's action means that Congress I wasted its time. Congress II, in turn, had other priorities that it considered more important, and is now being forced by the court to revisit the issue of endangered species protection rather than enact its own agenda. So the court ends up harming both the majority coalition that wanted the law enacted, and the majority

5:46 PM), http://www.slate.com/articles/news_and_politics/politics/2015/02/republicans_hope_for_the_supreme_court_to_kill_obamacare_if_the_court_rules.html [https://perma.cc/Y33K-JZ65].

297. Christiansen & Eskridge, *supra* note 195, at 1319; Richard L. Hasen, *End of the Dialogue? Political Polarization, the Supreme Court, and Congress*, 86 S. CAL. L. REV. 205, 209 (2013).

298. An established rule that courts must always choose a terrible remedy would also give the enacting Congress reason to clarify its remedial choice ex ante. But this would only work when Congress is able to anticipate a constitutional challenge. *See, e.g.*, Heckler v. Mathews, 465 U.S. 728 (1984).

coalition that later has to repeal the law.[299] In so doing, the court undermines the policy choices made by voting majorities in several democratic elections.

CONCLUSION

Jean-Paul Sartre writes of a waiter who really gets into his job. The waiter moves quickly and stiffly, bends toward customers a touch too eagerly, carries his tray with the recklessness of a tightrope walker.[300] This waiter is play-acting at being a waiter, and in so doing is engaged in a kind of illusory self-constraint. He acts as though he *is* a waiter, in an ultimate sense, as if that role encompassed the whole of his being. He denies, or perhaps forgets, that he has the freedom to do otherwise—that the role of "waiter" is merely a social artifact, a set of arbitrary customs connected to a trade. In fact, he freely chooses how he behaves while waiting tables, just as he freely chooses how he behaves in the rest of his life. But he has, for whatever reason, denied to himself that he exercises such choice.

The American doctrine of constitutional remedies embodies a similar illusory self-constraint. Judges sometimes pretend as though they face certain restrictions on account of their role. They cannot add language to a statute to make it constitutional. They cannot entertain a facial challenge when an as-applied remedy is available. They must adopt an avoidance reading rather than invalidate a statute. But none of these constraints actually exist—there is no deep justification for them, and judges commonly break them in practice. Rather than play-act at being unable to change the law, judges should acknowledge that they exercise substantial discretion when choosing constitutional remedies. This freedom may be unsought, but it does no good to deny its existence. Instead, American judges should elaborate an approach for exercising this freedom consistent with the logic of our system of government. The ambition of this Article has been to show that the best such approach is to find the remedy that the enacting legislature would have preferred.

299. This argument may have less force if the two congresses are the same, or relatively similar, in composition.

300. JEAN-PAUL SARTRE, BEING AND NOTHINGNESS: AN ESSAY ON PHENOMENOLOGICAL ONTOLOGY 59–60 (Hazel E. Barnes trans., 1956).