INTRODUCTION

A crucial part of the Supreme Court’s role is implementing the Constitution. To make the document into a set of legal rules that lower courts can enforce, the Court must translate the Constitution’s lofty values into effective rules. Because of this imperative, strategic concerns – choices about

---


2 See Amar, supra note 1, at 79-80; Fallon, supra note 1, at 56-57; Sager, supra note 1, at
how best to implement the Constitution – play an important role in shaping the
counters of virtually all constitutional doctrine.3

Perhaps the most prominent example is the Court’s crafting of prophylactic
rules to make constitutional rights “work better in practice.”4 Concerned that
constitutional violations might go undetected or unchecked under case-by-case
review, the Court regularly employs prophylactic rules to overcome this risk
and to better ensure the enforcement of constitutional rights.5 On the other
hand, the Court under-enforces certain constitutional rights, choosing, for
institutional reasons, a clear rule empowering the government to act rather than
a rule that enforces the Constitution to its full conceptual limits.6 Making
strategic judgments about how to implement constitutional rights, thus, is at the
very core of the Supreme Court’s function.

Few areas of law more sharply present the issue of the Court’s role in
implementing the Constitution than the law governing facial and as-applied
challenges to statutes. Term after Supreme Court term, many of the most hotly
contested constitutional questions turn on the choice between sustaining a
facial challenge or requiring affected parties to bring a series of as-applied
challenges. May a state enact a ban on partial-birth abortion?7 May a state
prohibit sodomy?8 May a school district authorize student-initiated prayer at football games?9 In each case, the Court’s invalidation depended on its
judgment that a facial challenge was the better means of implementing the
Constitution.

Facial and as-applied challenges present drastically different ways of
enforcing constitutional rights. As-applied adjudication proceeds deliberately,
case-by-case. Courts adjudicate rights-holders’ claims on an individualized
basis, and, over time, limit a statute’s reach by severing unconstitutional
applications.10 In a facial challenge, by contrast, the plaintiff challenges the
law in its entirety, seeking to enjoin the statute in all of its applications.11

961-63, 973.

3 See Evan H. Caminker, Miranda and Some Puzzles of “Prophylactic” Rules, 70 U.
    CIN. L. REV. 1, 25-26 (2001); Levinson, supra note 1, at 901; Sager, supra note 1, at 973;
    Strauss, supra note 1, at 195.

4 Levinson, supra note 1, at 866.

5 See id. at 899-904; Strauss, supra note 1, at 195.

6 See Lawrence G. Sager, Fair Measure: The Legal Status of Underenforced

    on its face).


    policy on its face).

10 For discussion of severability law, see generally John Copeland Nagle, Severability, 72
    N.C. L. REV. 203 (1993); Robert L. Stern, Separability and Separability Clauses in the
    Supreme Court, 51 HARV. L. REV. 76 (1937).

Facial invalidation renders the challenged law a legal nullity, and thus allows a court to vindicate the rights of many in a single litigation, making it a speedier and less costly remedy.

These differences have enormous practical significance. Take the case of a plaintiff challenging the validity of a restrictive abortion law. If the court allows a facial challenge, a single plaintiff can challenge the law on behalf of all women, and can stop the law from ever taking effect. If, however, the court finds a facial challenge improper, affected parties will have to bring as-applied challenges to vindicate their rights. Some women may not do so at all, and those who do may well receive a narrow ruling, perhaps even limited to their individual circumstances. Under the as-applied model, courts implement constitutional norms on a slower, more gradual basis.

In United States v. Salerno, two criminal defendants brought a facial challenge to the Bail Reform Act’s pretrial detention provisions on the grounds that due process did not permit detention based on a finding of future dangerousness. The Court rejected that argument, and, in doing so, prescribed a rule for adjudicating facial challenges: “A facial challenge to a legislative Act is the most difficult challenge to mount successfully, since the challenger must establish that no set of circumstances exists under which the Act would be valid.”

Salerno permits a facial challenge only in cases in which every plaintiff would win an as-applied challenge, i.e., those cases in which the government cannot validly apply the statute to anyone. Otherwise, to secure relief, a rights-holder must show that the law is unconstitutional as claim that a “statute is invalid in toto – and therefore incapable of any valid application”). Quite often, the plaintiff in a facial challenge is a rights-holder, but in some instances he or she need not be. See infra text accompanying notes 41-44; see infra notes 91, 105.

This is undoubtedly true for Supreme Court rulings, but less clear in the case of lower court rulings. Compare Richard Fallon, As-Applied and Facial Challenges and Third-Party Standing, 113 Harv. L. Rev. 1321, 1340 (2000) (arguing that a lower court ruling only applies to the particular plaintiffs in the case and is, at best, persuasive authority if the statute is applied to others) with Matthew Adler, Rights, Rules, and the Structure of Constitutional Adjudication: A Response to Professor Fallon, 113 Harv. L. Rev. 1371, 1411-12 (2000) (arguing that facial challenge doctrines must mark an exception from traditional law regarding claim preclusion against the government on which Fallon relies), and Matthew Adler, Rights Against Rules: The Moral Structure of American Constitutional Law, 97 Mich. L. Rev. 1, 149-50 (1998) [hereinafter Rights Against Rules] (same). In any event, as Adler notes, a plaintiff may bring a facial challenge as a class action and, if certified, can secure relief for all persons the law affects. Id. at 145-47.


Id. at 745. Salerno permits a single exception for First Amendment overbreadth claims. Id.
applied to him or her.\textsuperscript{16} 

Salerno aims to ensure that courts only provide relief to rights-holders, and does so by placing heavy reliance on the process of case-by-case adjudication. The doctrine is not concerned with the costs case-by-case adjudication might have on the enforcement of constitutional rights. Unless a statute is always unconstitutional, courts must use as-applied adjudication to determine the rights of constitutional claimants.\textsuperscript{17}

As-applied adjudication, however, is not cost-free. Its gradualism and more individualistic focus may, in certain circumstances, make more difficult the enforcement of constitutional rights. Salerno’s limitation is mistaken, I claim, precisely because it disregards these costs. A facial challenge may be appropriate because it is a better means of implementing the Constitution than requiring parties to mount a series of as-applied challenges because of the costs of case-by-case adjudication.\textsuperscript{18}

Neither courts nor scholars have recognized this second conceptual category

\textsuperscript{16} The Salerno rule, of late, is in retreat. In a number of recent cases, the Court has disregarded the Salerno rule and invalidated challenged statutes under a different, more lenient rule, without as much as a nod towards Salerno. See, e.g., Stenberg v. Carhart, 530 U.S. 914 (2000); Santa Fe Indep. Sch. Dist. v. Doe, 530 U.S. 290 (2000); Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833 (1992); see also Sabri v. United States, 124 S. Ct. 1941, 1948-49 (2004) (stating, in dicta, that overbreadth challenges are not restricted to First Amendment cases). In City of Chicago v. Morales, in which the Court invalidated a Chicago loitering ordinance as void for vagueness, a three-Judge plurality treated Salerno as empty dicta. 527 U.S. 41, 55 n.22 (1999). The dissenters, led by Justice Scalia, bitterly protested the Court’s facial invalidation and refusal to apply Salerno. Id. at 78-79 (Scalia, J., dissenting). Compare id. at 55 n.22 (plurality opinion) (refusing to apply Salerno) with id. at 74-82 (Scalia, J., dissenting) (demanding application of Salerno), and id. at 111 (Thomas, J., dissenting) (same). Despite these developments, Salerno still hangs on as official doctrine. As of yet, a majority of the Court has not repudiated or explicitly limited Salerno. Lower courts continue to apply Salerno, at least outside of abortion cases. See, e.g., Sherbrooke Turf, Inc. v. Minn. Dep’t of Transp., 345 F.3d 964, 971 (8th Cir. 2003); Hotel & Motel Ass’n of Oakland v. City of Oakland, 344 F.3d 959, 971-72 (9th Cir. 2003); Belitskus v. Pizzingrilli, 343 F.3d 632, 648 n.10 (3d Cir. 2003); Nebraska v. EPA, 331 F.3d 995, 998 (D.C. Cir. 2003); Doe v. Kearney, 329 F.3d 1286, 1294 (11th Cir. 2003); Doe v. Heck, 327 F.3d 492, 528 (7th Cir. 2003); Cranley v. Nat’l Life Ins. Co. of Vt., 318 F.3d 105, 110 (2d Cir. 2003). This Term, the Court will likely resolve whether Salerno applies in abortion cases. See Ayotte v. Planned Parenthood of N. New Eng., 125 S. Ct. 2294 (2005) (order granting certiorari).

\textsuperscript{17} In this regard, Salerno is an example of what Adrian Vermeule calls a maximizing rule of interpretation; its concern is to generate correct outcomes, regardless of cost. See Adrian Vermeule, Three Strategies of Interpretation, 42 SAN DIEGO L. REV. 607, 608-12 (2005).

\textsuperscript{18} In Vermeule’s terms, I claim that an optimizing rule – one that takes into account the costs of decisionmaking and interpretative error in order to optimize enforcement of constitutional rights – is more attractive than Salerno’s maximizing rule, even though, like all optimizing rules, it ends up protecting certain undeserving litigants. Id. at 613-16, 620 (critiquing maximizing rules for ignoring costs and suggesting that optimizing rules may often be superior).
of facial challenges, what I will call the strategic facial challenge. In a strategic challenge, a court must make a comparative judgment about the costs and benefits of facial invalidation vis-à-vis as-applied challenges. Facial invalidation depends on a court’s judgment that, due to a defect in the process of case-by-case adjudication, facial invalidation is a better means of implementing the Constitution than as-applied challenges. In this respect, a strategic facial challenge is a close relative of prophylactic rules. Like prophylactic rules, strategic facial challenges aim to better enforce constitutional rights by preempts case-by-case review because of the fear that such review will not adequately protect constitutional norms.

In fact, without naming it as such, the Court regularly employs this strategic device in a wide range of cases across many constitutional doctrines. From free speech and privacy cases to vagueness doctrine to Establishment Clause cases, the Court’s jurisprudence is replete with strategic facial invalidations. What unites these disparate cases is the Court’s conclusion that, in each of these areas of the law, facial invalidation is a better means of implementing the Constitution than case-by-case adjudication.

Emerging from this caselaw, I argue, are two rules that operationalize the conceptual category of strategic facial challenges. First, the challenged statute must infringe constitutional rights in a large or substantial number of cases. It is only in such circumstances that the benefits of facial invalidation outweigh its costs. Second, as-applied adjudication must not suffice to protect constitutional rights for some pragmatic reason. Facial invalidation is justified as a strategic tool to protect constitutional rights and guarantees – and a better means of implementing the Constitution – because as-applied adjudication is not an adequate protector of constitutional rights. In such circumstances, relying on case-by-case adjudication to narrow the law to its constitutional limits would, in of itself, cause constitutional harm. Facial invalidation prevents that harm, even though the result is to protect persons to whom the statute could be validly applied, persons who would lose an as-applied challenge to the law. Accordingly, a prerequisite for a strategic facial challenge is a showing that as-applied adjudication is inadequate to protect constitutional norms.

The archetype of a strategic facial challenge is a First Amendment overbreadth claim. Under that doctrine, a statute that prohibits a substantial amount of constitutionally protected speech is invalid on its face even though it has some valid applications. The doctrine’s purpose is to eliminate the chilling effect of a burdensome speech restriction. To leave such a statute in place and rely on case-by-case adjudication to separate valid from invalid

19 See infra Part II.B.
20 See infra Parts III-V.
22 Id. at 611-12.
applications would chill speech and lead to self-censorship. Because of the threat of chilled expression, facial invalidation is a better means of enforcing the First Amendment than case-by-case adjudication.

Overbreadth, however, is not a conceptual category in its own right; it is merely one kind of strategic facial challenge. Yet the literature on facial challenges is dominated by overbreadth doctrine and its chilling effect theory. Pages and pages have been written on the proper scope of overbreadth doctrine, without sufficient consideration to the other reasons for strategic facial invalidation. Scholars have focused on the overbreadth doctrine and chilling effects, and as a result, they have missed the variety of constitutional strategies inviting holdings of facial unconstitutionality.

In this Article, I argue that there are three strategic bases for facial invalidation that recur throughout constitutional adjudication, each corresponding to a distinct theory of why facial invalidation is a better means of implementing the Constitution than case-by-case adjudication. They are: (i) a chilling effect theory, featured not only in First Amendment overbreadth doctrine, but also in privacy, vagueness, voting, and travel cases; (ii) an excessive discretion theory that condemns statutes that confer too much discretion on actors to violate constitutional rights because (a) discriminatory abuse of that discretion might not be detected in case-by-case review and (b) case-by-case review of whether the actor abused that discretion might come too late to protect adequately constitutional rights; and (iii) a stigma theory, which calls for facial invalidation of laws that send a message of inequality because case-by-case review does not promise to eliminate expeditiously the law’s stigmatic message.

My argument proceeds as follows: Part I reviews the Salerno rule, explaining its roots in the idea that as-applied adjudication is the model of constitutional adjudication. Under Salerno, judges cannot facially invalidate a statute for strategic reasons, but only when the statute has no valid applications. Part II argues that the Salerno rule is too restrictive because it ignores the costs of as-applied adjudication and that courts should entertain facial challenges in cases in which they are a better means of implementing the Constitution than case-by-case adjudication. Overbreadth doctrine explicitly approves of such strategic invalidations, and the strategic protection of constitutional rights pervades constitutional jurisprudence. Accordingly,

23 See id. at 612.

24 See infra Part II.A.

25 See Adler, Rights Against Rules, supra note 12, at 161 (arguing that overbreadth doctrine is generally applicable to all kinds of constitutional challenges); Dorf, supra note 13, at 261-79 (arguing that overbreadth doctrine should apply to all constitutional rights that protect an individual’s right to engage in conduct free from government interference); Fallon, supra note 12, at 1352, 1354-55 (focusing on overbreadth’s chilling effect theory); Marc Isserles, Overcoming Overbreadth: Facial Challenges and the Valid Rule Requirement, 48 AM. U. L. REV. 359, 360-64 (1998) (defending Salerno’s treatment of overbreadth claims).
Strategic facial challenges are appropriate and reflect the Court’s regular creation of strategic doctrines throughout constitutional law. Parts III, IV, and V consider the question of when strategic facial invalidations are appropriate and, using the Court’s cases as a guide, argue that strategic invalidations are proper under three different theories: chilling effect, excessive discretion, and stigma.

I. Salerno’s Account of Constitutional Adjudication

The Salerno facial challenge rule rests on a narrow conception of federal judicial power. The Court’s role is not primarily to give meaning to or implement the Constitution, but to resolve what is essentially a private dispute between plaintiff and defendant. The idea – traceable in some form to Marbury v. Madison 26 – is that “[c]onstitutional judgments . . . are justified only out of the necessity of adjudicating rights in particular cases between the litigants brought before the Court.” 27 As-applied adjudication flows naturally from this conception of the judicial role. In as-applied adjudication, the Court’s constitutional pronouncements define the rights of litigants, and thus are a necessary part of resolving a controversy. 28 A facial challenge, on the other hand, goes beyond declaring the rights of the litigants; its very purpose is

---

26 5 U.S. (1 Cranch) 137 (1803). Marbury justified judicial review as a matter of necessity. To resolve the controversy between the parties, the Court had to decide which law to apply in cases of conflict between statute and higher law in the form of the Constitution. “[T]he very essence of judicial duty” consequently required the Court to decide whether the statute comport with the Constitution. Id. at 178. Thus, the Court assimilated constitutional adjudication to ordinary private law decisionmaking. See Henry P. Monaghan, Constitutional Adjudication: The Who and When, 82 YALE L.J. 1363, 1365-66 (1973). But Marbury itself violated this principle of necessity, making a number of important constitutional rulings that were unnecessary to resolve the dispute. See Richard H. Fallon, Marbury and the Constitutional Mind: A Bicentennial Essay on the Wages of Doctrinal Tension, 91 CAL. L. REV. 1, 15-16 (2003). As many scholars have written, a key part of Marbury recognizes the Court’s special function in ensuring adherence to the Constitution and justifies judicial review based on the need to ensure legislative obedience to the Constitution. See, e.g., id. at 14-16; Monaghan, supra, at 1370-71. For discussion of the tensions between Marbury’s private rights and special functions aspects, see Fallon, supra, at 33-36.

27 Broadrick v. Oklahoma, 413 U.S. 601, 611 (1973); see also United States v. Raines, 362 U.S. 17, 21 (1960) (“This Court . . . ‘has no jurisdiction to pronounce any statute . . . void, because irreconcilable with the Constitution, except as it is called upon to adjudge the legal rights of litigants in actual controversies.’” (quoting Liverpool, N.Y & Phila. S.S. Co. v. Comm’rs of Emigration, 113 U.S. 33, 39 (1885))).

28 See Massachusetts v. Mellon, 262 U.S. 447, 488 (1923) (“[T]he power exercised is that of ascertaining and declaring the law applicable to the controversy. It amounts to little more than the negative power to disregard an unconstitutional enactment, which otherwise would stand in the way of the enforcement of a legal right.”). Implicit here, of course, is the requirement that the litigants themselves have standing to sue, and thus that a controversy within the meaning of Article III is present.
to obtain a judicial declaration that the statute is invalid as a whole and cannot be enforced against anyone.

The *Salerno* facial challenge rule flows from the as-applied model of constitutional adjudication. It permits facial invalidation solely as an outgrowth of the rights-declaring function of case-by-case adjudication, i.e., when a declaration that the statute cannot be applied to the plaintiff “unmistakably . . . yield[s] the conclusion that a statute is invalid, not merely as applied to the facts, but . . . in whole.”29 The necessary result of the court’s ruling in a particular plaintiff’s case is that the plaintiff and every other person affected by the law could bring a successful as-applied challenge. Because every party subject to the statute would obtain as-applied relief, the court will invalidate the statute on its face.30 Thus, facial relief promotes judicial economy; rather than require every affected party to challenge the statute as-applied, *Salerno* permits all the as-applied claims to be bundled into a single facial challenge, thus vindicating each person’s constitutional rights.31

What courts cannot do, *Salerno* says, is act as risk managers, using facial invalidation strategically to guard against the risk that constitutional rights will be lost during the process of case-by-case adjudication. The *Salerno* rule is not at all concerned with the fact that case-by-case adjudication may impose costs on rights-holders. Its concern is to correctly resolve constitutional disputes,

---

29 Fallon, *supra* note 12, at 1337. Short of these circumstances, *Salerno* says, facial invalidation is unwarranted because courts can excise unconstitutional applications over time using severability doctrine. See *Dorf*, *supra* note 13, at 250; Fallon, *supra* note 12, at 1342.

30 See *Adler, Rights Against Rules*, supra note 12, at 154-55. As Adler explains: *Salerno* says . . . the following: Given some rule *R*, a court should facially invalidate *R* only if, for every person *X* against whom *R* might be enforced, the application of *R* to *X* would be unconstitutional. . . . *Salerno* starts with the notion that constitutional adjudication is, centrally, as-applied adjudication. What concerns the reviewing court, first and foremost, is whether *X*’s own treatment is unconstitutional.

*Id.*

31 The *Salerno* rule, at the same time, is in serious tension with the as-applied model of adjudication: it permits facial invalidation even though such relief is wholly unnecessary to declare the rights of the parties before the court. As Justice Scalia candidly observed, [i]t seems . . . fundamentally incompatible with [as-applied adjudication] for the Court not to be content to find that a statute is unconstitutional as applied to the person before it, but to go further and pronounce that the statute is unconstitutional in all applications. Its reasoning may well suggest as much, but to pronounce a holding on that point seems to me no more than an advisory opinion – which a federal court should never issue at all, and especially should not issue with regard to a constitutional question, as to which we seek to avoid even nonadvisory opinions.

City of Chicago v. Morales, 527 U.S. 41, 77 (1999) (Scalia, J., dissenting) (citations omitted). *Salerno*, therefore, ignores the fact that the as-applied model, taken to its logical conclusion, would not permit facial challenges at all. Put another way, *Salerno* adopts the as-applied model in weak form. The strong form of the as-applied argument would bar all facial challenges.
regardless of cost. Salerno, for that reason, prohibits courts from invalidating a statute on its face because invalidation is a better means of implementing the Constitution than case-by-case adjudication.32

Strategic facial invalidations do not strictly focus on protecting the plaintiff’s rights, and that is why Salerno rejects them. In a strategic facial challenge, the court’s concern is how – for the entire class subject to the statute – to best implement the Constitution. For that reason, the strategic challenge looks chiefly to the parties not before the court, and justifies facial invalidation on instrumental grounds, namely the inadequacies of the case-by-case process. Strategic facial invalidation is not an outgrowth of declaring the plaintiff’s rights; its very purpose is to protect the rights of all subject to the law, particularly absent parties who might lose their rights under a case-by-case regime. This very feature is what renders the strategic invalidation illegitimate in Salerno’s eyes. It is this feature that transforms courts from adjudicative bodies into “roving commissions”33 that supervise lawmaking, and results in excessive judicial intrusion into the work of the legislative and executive branches.34

II. SALERNO’S PROBLEM: STRATEGIC CONSTITUTIONAL RULEMAKING

The Salerno rule cannot withstand scrutiny. First Amendment overbreadth doctrine rightly recognizes that a strategic facial challenge is a legitimate tool to protect constitutional rights. Moreover, functionally identical strategic

32 See Morales, 527 U.S. at 78 (Scalia, J., dissenting) (“[W]e should not be holding a statute void in all its applications unless it is unconstitutional in all its applications.”).
34 What about standing rules prohibiting third-party claims? Isn’t Salerno simply a reflection of this settled standing principle? This argument fails. The third-party standing ban is not part of Article III’s constitutional requirements, but is a rule of prudence only, and thus is subject to modification by the Court as strategic concerns require. See Barrows v. Jackson, 346 U.S. 249, 255-57 (1953); Note, Standing To Assert Constitutional Just Tertii, 88 HARV. L. REV. 423, 424-25 (1974). Third-party standing doctrine allows a plaintiff to raise the rights of a third party and invalidate a statute, even though the statute does not violate the plaintiff’s own rights. Thus, from the standpoint of Article III, it is perfectly proper for a court to adjudicate a facial challenge to a statute raised by a party whose own rights are not in question. See Adler, Rights Against Rules, supra note 12, at 133-45. So understood, third-party standing law cannot possibly justify the Salerno rule. Third-party standing law makes room for the kinds of strategic rulemaking that Salerno forecloses. What is more, Salerno, no less than strategic challenges, allows a party to seek relief not only for himself or herself but for the entire class which the law burdens. If this is a third-party standing problem, it afflicts Salerno and strategic challenges equally. Disagreement over the Salerno rule is not a disagreement over the ‘standing’ question whether the person challenging the statute can raise the rights of third parties: under both Salerno and [a strategic facial challenge] he can. The disagreement relates to how many third-party rights he must prove to be infringed by the statute before he can win . . . . That is not a question of standing but of substantive law.
Morales, 527 U.S. at 80 n.3 (Scalia, J., dissenting).
decisionmaking is a regular feature of constitutional rulemaking. Throughout constitutional law, courts craft strategic doctrines to optimize the protection of constitutional rights, and, in doing so, often “overprotect” rights, preempting more individualized case-by-case review to better ensure the enforcement of constitutional rights. Optimizing rules pervade constitutional jurisprudence. Strategic facial challenges fit comfortably within this tradition of strategic rulemaking.

A. First Amendment Overbreadth Doctrine

Overbreadth presents a radically different account of the judicial role in adjudication. It rejects the idea that facial invalidation can only be an outgrowth of case-by-case adjudication. Overbreadth, quite explicitly, displaces as-applied adjudication and licenses the strategic use of facial invalidation because the as-applied method does not adequately protect constitutional rights. Far from converting courts into roving super-legislatures, overbreadth sees strategic invalidation as a necessary part of constitutional adjudication, properly employed when facial invalidation is a better means of enforcing constitutional rights than as-applied challenges.35

Overbreadth’s fear is that sweeping laws targeting speech – by their very existence – will chill persons from exercising their rights out of fear of state sanctions36 and that case-by-case remedies will be ineffective37. As-applied relief may vindicate the rights of the plaintiff, but it will provide no assurances and no protection against chilling effects to third-party speakers. Facial invalidation is needed to protect these absent third parties.

First Amendment overbreadth doctrine arose, beginning in the 1940s,
against the backdrop of labor and civil rights activism.\textsuperscript{38} In a series of cases beginning with *Thornhill v. Alabama*,\textsuperscript{39} the Court confronted sweeping speech restrictions that, on their face, barred a wide range of protected speech, and could easily be used by the state to censor and jail protestors. The combination of vulnerable speakers, powerful state criminal sanctions, and the Supreme Court’s limited resources led to the creation of the doctrine. Recognizing the powerful weapon the criminal statutes gave to prosecutors, the Court recognized that speakers would be protected weakly, at best, through case-by-case adjudication. Although some speakers might be ultimately vindicated, many would silence themselves rather than running the risk of prosecution and would not bring suit. To overcome this problem, the Court turned to facial invalidation.\textsuperscript{40}

Aiming to combat chilling effects, First Amendment overbreadth doctrine departs from the case-by-case framework and facilitates facial invalidation in two respects. First, it provides an across-the-board, automatic exception to the prohibition on third-party standing,\textsuperscript{41} thereby “allowing the first party with Article III standing to prevail, regardless of the character of his own activity.”\textsuperscript{42} As a result, any speaker subject to sanctions under the law can bring a facial challenge, even if the plaintiff has not engaged in protected speech and is forced to rely on the rights of others to invalidate the statute.\textsuperscript{43} Courts thus need not wait for the statute’s first application to constitutionally protected speech to determine its facial validity. Courts will enjoin an overbroad law at the behest of the first challenger.\textsuperscript{44}

\textsuperscript{38} See Fallon, * supra* note 36, at 863 (discussing the historical background of overbreadth doctrine).

\textsuperscript{39} 310 U.S. 88 (1940).


\textsuperscript{41} See, e.g., *Broadrick*, 413 U.S. at 612.

\textsuperscript{42} See *Note*, * supra* note 13, at 856.

\textsuperscript{43} See Alexander v. United States, 509 U.S. 544, 555 (1993) (“The ‘overbreadth’ doctrine, which is a departure from traditional rules of standing, permits a defendant to make a facial challenge to an overly broad statute restricting speech, even if he himself has engaged in speech that could be regulated under a more narrowly drawn statute.”); cf. *Yazoo & Miss. Valley R.R. Co. v. Jackson Vinegar Co.*, 226 U.S. 217, 219-20 (1912) (rejecting facial challenge under traditional standing rule because plaintiff was asserting rights of absent third parties, and upholding validity of statute as applied to plaintiff).

\textsuperscript{44} Some of the Court’s cases suggest that the standing component of overbreadth is a double-edged sword, meaning that if the plaintiff’s speech is constitutionally protected, a court should consider plaintiff’s claim only as applied. See Colo. Republican Campaign Comm. v. *FEC*, 518 U.S. 604, 623-24 (1996) (plurality opinion); United States v. Nat’l Treasury Employees Union, 513 U.S. 454, 478 (1995); *id.* at 486 (O’Connor, J., concurring in part and dissenting in part); *Brockett v. Spokane Arcades*, Inc., 472 U.S. 491, 504 (1985). But these statements must be qualified in light of the long line of cases in which the Court has facially invalidated statutes under overbreadth doctrine at the behest of protected
Second, the doctrine mandates the facial invalidation of substantially overbroad laws. Courts may not leave in place sweeping speech restrictions because they have some valid applications, and may not rely on case-by-case adjudication and severability doctrine to narrow an overbroad law over time. To eliminate the chilling effect that the statute’s very existence produces, overbreadth renders a substantially overbroad statute void as a whole, barring the government from enforcing the law at all, even in its constitutional applications. This forces the legislature back to the drafting table should it wish to regulate the expressive activities and creates a salutary incentive for legislators to write narrow statutes when regulating free speech.

The notion of substantial overbreadth is directly related to the doctrine’s strategic goal of eliminating a chilling effect. Only in the case of substantially overbroad laws – laws with “the potential to repeatedly chill the exercise of expressive activity by many individuals” – does the strategic goal of preventing a chilling effect justify facial invalidation. As Justice Scalia’s recent opinion in *Hicks* shows, the substantial overbreadth inquiry calls for a strategic judgment about the costs and benefits of facial versus as-applied review. The requirement of substantial overbreadth limits facial invalidation

---

45 See *Ashcroft v. Free Speech Coal.*, 535 U.S. 234, 244 (2002); *Reno v. ACLU*, 521 U.S. 844, 882-84 (1997); Bd. of Airport Comm’rs v. Jews For Jesus, Inc., 482 U.S. 569, 574-77 (1987); *see also* Watchtower Bible & Tract Soc’y of N.Y., Inc. v. Vill. of Stratton, 536 U.S. 150, 166 n.14, 167 (2002) (adjudicating hybrid first/third-party overbreadth claim). These cases suggest that a protected speaker should be able to sustain an overbreadth challenge where he or she shows that the statute significantly compromises the First Amendment rights of third-party speakers.


49 Note, supra note 13, at 859 (“[O]nly substantially overbroad laws set up the kind and degree of chill that is judicially cognizable.”). Substantial overbreadth has both a quantitative and a qualitative component. Overbreadth must be substantial “not only in an absolute sense, but also relative to the scope of the law’s plainly legitimate applications.” *Hicks*, 539 U.S. at 120. The doctrine thus asks whether the statute would be invalid in most or many of the cases contemplated by the statute. See *Fallon*, supra note 36, at 894. At the same time, this is not solely a question of numbers; inevitably, courts have to make choices about when First Amendment rights require the stronger protection overbreadth affords. See *id.* at 894-97.

50 See *Hicks*, 539 U.S. at 118-20.
to those cases in which the doctrine’s benefits – protecting a significant number of speakers against chilling effects and forcing legislatures to regulate speech with precision – outweigh the social costs of forbidding all enforcement of a statute that can be validly applied in some cases.\textsuperscript{51}

Under the doctrine, the costs and benefits tip in favor of facial invalidation of substantially overbroad laws for three reasons. First, a substantially overbroad law chills expressive activities, and only facial invalidation will eliminate the chilling effect.\textsuperscript{52} Second, where a statute is substantially overbroad, the costs of facial invalidation are low. A substantially overbroad law, by definition, impermissibly intrudes on free speech rights in many or most of the cases contemplated by the statute, and would be so declared if plaintiffs brought as-applied challenges.\textsuperscript{53} Facial invalidation, by and large, thus protects deserving speakers – individuals who would engage in speech that the Constitution protects but might censor themselves for fear of sanctions in the absence of facial invalidation. Only in a small number of cases does facial invalidation protect conduct that the legislature might legitimately proscribe. That means that the benefits of invalidation are significant and the costs are low. Third, facial invalidation better respects the separation of powers between courts and legislatures. To cure the defects in a substantially overbroad statute almost always requires a court to perform radical surgery on the statute and threatens to enmesh the court in policy choices that legislatures are better suited to make.\textsuperscript{54} Facial invalidation avoids these costs, forcing the legislature to revisit its regulation, and, if it so desires, enact a narrower statute that more precisely targets conduct that it can rightfully proscribe.

Combating chilling effects is the primary, but not the sole, concern of First Amendment overbreadth doctrine. Free speech licensing cases feature an additional strategic justification for invalidation.\textsuperscript{55} A licensing scheme that gives too much discretion to the licensor not only chills speakers, but is also a tool for viewpoint discrimination; it “has the potential for becoming a means of suppressing a particular point of view.”\textsuperscript{56} In the same way that as-applied

\textsuperscript{51} See id. at 119.

\textsuperscript{52} See Note, supra note 13, at 852-58 (arguing that only facial invalidation is sufficient to protect against the chilling effect of overbroad statutes).

\textsuperscript{53} See supra note 49.

\textsuperscript{54} See Reno v. ACLU, 521 U.S. 844, 884-85 (1997); Dorf, supra note 13, at 292-93.

\textsuperscript{55} Overbreadth condemns licensing statutes that “vest unbridled discretion in a government official,” City of Lakewood v. Plain Dealer Publ’g Co., 486 U.S. 750, 755 (1988), because the law “creates an impermissible risk of suppression of ideas.” Forsyth County v. Nationalist Movement, 505 U.S. 123, 129 (1992). Because of this risk, the Court calls for facial invalidation “even though [the law’s] application in the case under consideration may be constitutionally unobjectionable.” Nationalist Movement, 505 U.S. at 129.

\textsuperscript{56} Id. at 130 (quoting Heffron v. Int’l Soc’y for Krishna Consciousness, 452 U.S. 640, 649 (1981)); see also Thomas v. Chi. Park Dist., 534 U.S. 316, 323 (2002) (“Where the licensing official enjoys unduly broad discretion in determining whether to grant or deny a
litigation cannot cure the chilling effects of an overbroad law, it is a toothless remedy against the possibility of discrimination arising from a discretionary licensing scheme. As applied, it is extremely difficult for a court to assess whether a license denial was made because of the content of the plaintiff’s speech, when there are no clear and objective standards against which to measure the denial. Therefore, under case-by-case review, some constitutional violations will go unchecked because content discrimination will pass under the judicial radar. What is more, mounting an as-applied claim of discrimination is an onerous, time-consuming process, during which free expression remains suppressed. Facial invalidation is thus a tactic to prevent both chilling effects and content and viewpoint discrimination in the face of the defects of as-applied adjudication.

The account sketched above assumes that courts can accurately gauge whether a statute’s overbreadth is substantial or not. But critics of overbreadth argue that courts are not fit for the task and that, as a result, they invalidate statutes when they should not. If that were correct, the error costs might counsel against invalidation and in favor of a case-by-case process in which courts deal only with the parties before them.

Overbreadth, of course, creates a risk that courts will erroneously invalidate statutes. Analyzing a statute as a whole is a more difficult task than simply analyzing a single application, so the risk of judicial error is higher. But there is no reason to believe these error costs should lead us to question overbreadth’s demand for facial invalidation. First, overbreadth challengers bear the burden of proof, so doubts are resolved against invalidation. In close cases, a court can reject the facial challenge if unconvinced that the statute’s overbreadth is truly substantial. Second, the likelihood that courts will permit, there is a risk that he will favor or disfavor speech based on its content.”).
systematically err in analyzing “substantial overbreadth” is scant at best. Overbreadth does not call for complicated inquiries; it simply demands that the court interpret the statutory text to understand its potential scope and assess whether a statute unconstitutionally intrudes on protected speech in a great many cases or not. This task is well within the judicial ken and only marginally different from the assessment Salerno requires. Third, overbreadth helps prevent judicial errors that might arise from case-by-case adjudication. As the licensing cases make clear, overbreadth demands facial invalidation because courts will inevitably make errors if they try to assess whether a license denial reflects illicit content discrimination. In these cases, facial invalidation avoids – not causes – errors. Therefore, used properly, the benefits of facial invalidation outweigh any costs resulting from judicial error.

B. **Strategic Overprotection in Constitutional Rulemaking**

Salerno’s answer to First Amendment overbreadth doctrine is to marginalize it. Salerno concedes its validity, but labels it the exception to the normal rules governing facial challenges. Salerno seeks to cabin overbreadth (and strategic invalidations more generally) by conceptualizing the doctrine as a special creature of the First Amendment. Salerno presumes that overbreadth’s strategic overprotection of constitutional rights is exceptional and must be strictly confined. In other words, the favored status of First Amendment rights justifies giving courts a remedial power they otherwise lack. In all other cases, Salerno says, strategic facial challenges are inappropriate. The Court’s role is to resolve disputes brought by specific aggrieved individuals; therefore, the judicial focus must be on whether the statute is unconstitutional as applied to the party challenging the legislature’s act. As Salerno provides, this leaves room for a facial challenge only in those cases in which a statute can never be validly applied, that is, in cases in which every as-applied claim would succeed.

The difficulty with this argument is that it assumes that the Constitution establishes as-applied adjudication as the baseline against which constitutional claims should be measured. The Constitution says little, if anything, about judicial review. It says even less about how judicial review should operate or how the Supreme Court should enforce the protections spelled out in the Constitution. The Constitution sets forth many specific guarantees of rights, but provides little guidance on how courts should enforce these guarantees. The Supreme Court’s role is to fill these gaps, working much like a common

---

62 See supra text accompanying notes 55-59.
64 Id.
65 See Fallon, supra note 36, at 884 (“First Amendment rights are . . . special, and any chill of their exercise gives rise to extraordinary constitutional concerns.”); Note, supra note 13, at 852 (arguing that overbreadth doctrine “find[s] justification in the favored status of rights to expression and association in the constitutional scheme”).
law court in translating the document’s guarantees into enforceable legal rules. In doing so, the Court must weigh the costs and benefits of various legal regimes in the search for the doctrine that best enforces the underlying constitutional right. This means that the Court will have to assess which mechanism of judicial review – a facial challenge or an as-applied challenge – is proper for adjudicating the constitutional challenge.

As-applied adjudication, of course, carries with it important benefits. It conserves judicial firepower, forcing courts to proceed slowly, hearing one case at a time, and deciding that case on the facts generated by that concrete controversy. This not only limits judicial intervention into the work of other branches of government, it ensures that courts do not make uncertain speculations about how a law operates outside of the facts generated by the controversy before it. But as-applied adjudication also imposes serious costs. Proceeding one case at a time is a slow process that may make it more difficult to enforce the rights of persons not before the court. A court concerned only with resolving the specific case before it may be blind to this systemic harm to constitutional rights. As-applied adjudication may also have a distorting effect. A focus on the case before it might cause a court to lose sight of a larger constitutional injury. A facial challenge – because it forces a court to look through a wider lens – may allow a court to avoid these costs.

Salerno, however, says we need not balance these costs and benefits in deciding how to best enforce constitutional rights; indeed, it claims that we need not pay any attention to the costs of case-by-case adjudication at all. But Salerno ignores the fact that courts, throughout constitutional jurisprudence, craft strategic doctrines that preempt case-by-case review in an effort to make constitutional rights work better in practice, and this proves its undoing. To understand why, we need to look more carefully at a foundational example of such strategic rulemaking: Miranda v. Arizona.

The Miranda Court formulated its famous warnings requirement for two reasons. First, the Court found a substantial risk that police officers would

---

66 See Fallon, supra note 1, at 56-57 (discussing the Court’s role in implementing the Constitution); David Strauss, Common Law Constitutional Interpretation, 63 U. CHI. L. REV. 877 (1996) (emphasizing commonalities between constitutional and common law adjudication).

67 See David A. Strauss, Miranda, the Constitution, and Congress, 99 Mich. L. Rev. 958, 962-63 (2001) (“The crucial point . . . is that every principle of constitutional law reflects, implicitly, a . . . balancing of costs and benefits. No principle enforces itself; no principle can be perfectly administered.”).

68 See Cass R. Sunstein, One Case at a Time: Judicial Minimalism on the Supreme Court 3-24 (Harv. Univ. Press 1999) (describing the benefits of judicial minimalism, i.e., deciding only the specific controversy before the Court).


70 See supra text accompanying note 32.

violate Fifth Amendment rights when they interrogated suspects in the coercive atmosphere of the station house.\textsuperscript{72} Second, the Court concluded that case-by-case review of confessions for voluntariness was unworkable and inadequate to protect Fifth Amendment rights.\textsuperscript{73} The Court’s solution to the substantial risk of Fifth Amendment violations was to mandate warnings — although the warning regime overprotected the Fifth Amendment self-incrimination guarantee by excluding unwarned statements by suspects who knew, but chose not to exercise their rights\textsuperscript{74} — because of the defects of case-by-case voluntariness review.\textsuperscript{75}

On first blush, \textit{Miranda} seems irrelevant to the law of facial challenges as it did not even involve a constitutional challenge to a statute. Yet the two doctrines are, in fact, closely related. \textit{Miranda} features a judgment about how to best enforce constitutional rights indistinguishable from strategic facial invalidation: overprotection in favor of case-by-case review. Strategic facial invalidation is a judicial response to the inadequacy of case-by-case adjudication to protect constitutional rights; the \textit{Miranda} rule is a response to the inadequacy of case-by-case voluntariness review. The case-by-case method used for decades prior to \textit{Miranda} was too porous a sieve to protect Fifth Amendment rights. It produced too many errors, due to both the difficulty of ascertaining later what occurred during an interrogation and the

\textsuperscript{72} Dickerson v. United States, 530 U.S. 428, 435 (2000) (“[T]he coercion inherent in custodial interrogation blurs the line between voluntary and involuntary statements, and thus heightens the risk that an individual will not be ‘accorded his privilege under the Fifth Amendment . . . not to be compelled to incriminate himself.’” (quoting \textit{Miranda}, 384 U.S. at 439)); Moran v. Burbine, 475 U.S. 412, 426 (1986) (“[T]he interrogation process is ‘inherently’ coercive and . . . , as a consequence, there exists a substantial risk that the police will inadvertently traverse the fine line between legitimate efforts to elicit admissions and constitutionally impermissible compulsion.”).

\textsuperscript{73} See Missouri v. Seibert, 124 S. Ct. 2601, 2607-08 (2004) (plurality opinion) (arguing that the voluntariness inquiry “posed an ‘unacceptably great’ risk that involuntary custodial confessions would escape detection” (quoting \textit{Dickerson}, 530 U.S. at 444)); see also Haley v. Ohio, 332 U.S. 596, 606 (1948) (Frankfurter, J., concurring) (“Unhappily we have neither the physical nor intellectual weights and measures by which judicial judgment can determine when pressures in securing a confession reach the coercive intensity that calls for the exclusion of a statement so secured.”).

\textsuperscript{74} Some scholars have defended \textit{Miranda} as nothing more than an exercise of the Court’s traditional role in constitutional interpretation. Michael Dorf and Barry Friedman, for example, argue that \textit{Miranda} decided that the Fifth Amendment requires that the accused learn of the right not to speak with the police, and this understanding of the Fifth Amendment justifies the warning regime. \textit{See} Michael C. Dorf & Barry Friedman, \textit{Shared Constitutional Interpretation}, 2000 SUP. CT. REV. 61, 76-79. The difficulty with this claim is that it does not explain why \textit{Miranda} excludes unwarned, but voluntary, confessions by suspects who knew their rights — confessions that can hardly be described as compelled within the meaning of the self-incrimination guarantee.

\textsuperscript{75} For useful discussions of \textit{Miranda}, see Berman, supra note 1, at 114-32; Caminker, supra note 3, at 9-11; Strauss, supra note 67, at 961-66; Strauss, supra note 1, at 195-209.
vagueness of the voluntariness standard itself. To overcome these problems, the Court put in place an overprotective, easy-to-apply rule. Even though some undeserving defendants might go free because *Miranda* excludes their unwarned voluntary confessions, *Miranda* better enforces the Fifth Amendment privilege, thereby justifying such overprotection. A distrust of case-by-case review unifies the two doctrines. In both, the Court imposes an overprotective rule – in overbreadth, facial invalidation; in *Miranda*, the warning requirement – rather than using case-by-case analysis to eliminate the possible loss of constitutional rights. Both are strategic doctrines designed for the purposes of better enforcing constitutional rights.76

The same types of strategic concerns underlying *Miranda* occupy much of constitutional law. Because courts need to make the Constitution into legally enforceable doctrine, they must constantly make strategic judgments about how to protect a given right, and, quite often, preempt case-by-case inquiries with broad rules because of the defects of case-by-case review. As David Strauss and a host of other scholars have shown, numerous constitutional doctrines – not simply in First Amendment cases, but across constitutional jurisprudence – impose a prophylactic rule at the expense of case-by-case review, reflecting the Court’s strategic judgment that a strict categorical rule is a better way to enforce the constitutional right than a case-by-case approach.77 Where a case-by-case approach is too weak or otherwise inadequate to protect the constitutional guarantee, the Court has no compunction in setting the case-by-case process aside, as it did in *Miranda*.78

76 See Strauss, *supra* note 1, at 197 (drawing a parallel between overbreadth licensing cases and *Miranda*). Levinson sees a similar connection between First Amendment overbreadth doctrine and *Miranda*:

Declaring licensing schemes facially unconstitutional...is best understood...as enforcing a prophylactic rule – just like the one in *Miranda*. The rule is designed to prevent certain types of content or viewpoint discrimination that offend free speech principles, but in order to do so efficiently, it must forbid some government activity that is consistent with the ‘real’ First Amendment (i.e. open-ended licensing schemes in which discretion is not abused).

Levinson, *supra* note 1, at 902.

77 These rules include a host of constitutional criminal procedure rules, the one-person, one-vote rule, and rules against content-based and race-based discrimination, among others. See, e.g., Susan R. Klein, *Identifying and (Re)formulating Prophylactic Rules, Safe Harbors, and Incidental Rights in Constitutional Criminal Procedure*, 99 Mich. L. Rev. 1030, 1037-44 (2001) (discussing pervasive use of prophylactic rules in constitutional criminal procedure doctrines); Levinson, *supra* note 1, at 887, 899-904 (discussing the prophylactic aspects of the-one person, one-vote rule); Strauss, *supra* note 1, at 198-205 (discussing prophylactic character of much First Amendment jurisprudence).

78 The necessity of these strategic tools is, in part, a consequence of the Court’s limited resources. The Court can only hear a tiny fraction of constitutional disputes, and, inevitably, must consider whether a relevant right should be protected through case-by-case review or a broader, overprotective rule. It cannot – as the pre-*Miranda* voluntariness cases bear out – commit so much of its precious resources to norm-application as opposed to norm creation.
Strategic protection also runs through the law of constitutional remedies. In criminal procedure cases, to choose one of many examples, the Court regularly mandates strategic remedies to better protect constitutional rights. The most famous of these is the Fourth Amendment’s exclusionary rule, which has long been justified as a deterrent remedy. Exclusion is the appropriate constitutional remedy because it deters police officers from conducting unconstitutional searches and seizures. Another important example is the rule that a defendant has the right to overturn his or her conviction upon a showing of racial discrimination in grand jury selection, notwithstanding that the defendant was convicted after a fair trial at which discrimination did not occur. The Court mandates reversal – just as it mandates exclusion in Fourth Amendment cases – to deter unconstitutional state action; alternative remedies, while theoretically available, are likely to be ineffective, and, as a result, the Court put in place a strategic remedy to deter unconstitutional behavior. Other examples abound. The Court has often put its imprimatur on injunctions that not only require the defendant to obey the law, but also mandate that the defendant take certain steps (not otherwise required by law) to ensure that threatened violations do not recur. These injunctions are strategic devices that better ensure the protection of constitutional rights.

In showing that prophylactic rules pervade constitutional law, Strauss’s main concern was to defend *Miranda* against attacks that it was illegitimate, but his argument has important, and so far unrecognized, implications for how we assess facial challenges. *Salerno*’s “all circumstances” rule says that courts should not invalidate statutes for strategic reasons, but only when every affected person would win an as-applied challenge, save in exceptional circumstances, i.e., First Amendment overbreadth cases. Strauss’s claim powerfully undermines the *Salerno* rule. If courts routinely develop prophylactic rules to better protect constitutional rights, strategic facial invalidation should not be limited to First Amendment overbreadth cases. As a general matter, courts should permit strategic challenges whenever facial invalidation is a better means of enforcing constitutional rights because of a

---


80 Id.


82 See Meltzer, *supra* note 79, at 253-78.

Making room for strategic invalidations does not require ousting the Salerno rule so much as reformulating it. Salerno should be regarded as an important default rule. It holds that, absent a strategic reason for invalidation, courts should not invalidate statutes on their face if the statutes have valid applications. There are substantial justifications for treating Salerno as a default rule. If a statute has valid applications and no harm occurs in using case-by-case adjudication, facial invalidation seems gratuitous. These policies, however, must give way when invalidation is necessary, for strategic reasons, to protect constitutional rights. Given the pervasiveness of strategic constitutional rulemaking, courts should have the authority to invalidate statutes on their face for strategic reasons.

III. THE CHILLING EFFECT STRATEGY

If strategic facial challenges are proper exercises of judicial authority, in which cases should courts utilize them? The Supreme Court’s jurisprudence, we shall see, tells a very different story than Salerno does. It illustrates, in line with my claim, that strategic invalidations are part and parcel of judicial remedies in constitutional cases, used to protect constitutional rights when as-applied adjudication does not suffice.

Two legal requirements for strategic facial invalidation emerge from the cases. First, the challenged statute must violate constitutional rights in a substantial number of cases. Second, case-by-case adjudication must be inadequate to protect constitutional rights. These two requirements reflect the cost-benefit analysis that we observed in First Amendment overbreadth doctrine: facial invalidation is proper because (i) it avoids a defect in case-by-case adjudication that would impede enforcement of the constitutional right in question; (ii) its benefits are high and its costs low when a statute violates constitutional rights in a substantial number of cases; and (iii) it respects the separation of powers between courts and legislatures by avoiding the kind of radical surgery the court would need to perform to cure the statute’s substantial defects. In the remainder of this Article, I argue that there are three

84 The Court properly applied this default rule in Salerno. Because the class subject to the pretrial detention provision had been haled into court, and had an obvious opportunity to challenge their detentions, the default rule applied. For other cases applying this default rule, see, e.g., Reno v. Flores, 507 U.S. 292 (1993) (rejecting facial challenge to provision authorizing detention of certain juvenile aliens); Schall v. Martin, 467 U.S. 253 (1984) (rejecting facial challenge to provision authorizing juvenile detention); County Court v. Allen, 442 U.S. 140 (1979) (rejecting facial challenge to statute allowing permissive inference of gun ownership in criminal cases); United States v. Nat’l Dairy Prods., Inc., 372 U.S. 29 (1963) (rejecting vagueness facial challenge to antitrust laws); Yazoo & Miss. Valley R.R. Co. v. Jackson Vinegar Co., 226 U.S. 217 (1912) (rejecting facial challenge to economic regulation of railroads). What unites all of these cases is the absence of any strategic reason for invalidation.

85 See supra text accompanying notes 51-54.
invalidation strategies, each resting on a different reason to favor facial invalidation over as-applied challenges as a better means of implementing the Constitution.

The paradigmatic, and dominant, strategy is the chilling effect strategy, which has been featured in First Amendment overbreadth cases. But the chilling effect move is not unique to First Amendment cases; it logically applies to all constitutional protections of an individual’s right to engage in conduct. Indeed, across constitutional law, the Court has used a chilling effect narrative to justify facial invalidation of laws that burden constitutionally protected rights. This narrative connects First Amendment overbreadth cases with abortion cases, fundamental rights vagueness cases, right to vote cases, and right to travel cases.

A. The Right to Privacy

In Planned Parenthood of Southeastern Pennsylvania v. Casey, the Court applied the functional equivalent of First Amendment overbreadth doctrine to invalidate Pennsylvania’s husband notification requirement. The Court invalidated the provision because, given the realities of domestic violence, forced notification was “likely to prevent a significant number of women from obtaining an abortion.” Because it would deter many women from seeking abortions, the law was invalid on its face; it did not matter that the law might not impose an unconstitutional burden in all instances. As the joint opinion explained, the statute was invalid on its face because “in a large fraction of the cases in which [the statute] is relevant, it will operate as a substantial obstacle to a woman’s choice to undergo an abortion.”

---

86 See Dorf, supra note 13, at 265 (arguing that “the kind of chilling effect that justifies the First Amendment’s overbreadth doctrine” also applies to any other “right to engage in primary conduct”); Fallon, supra note 36, at 861 n.48 (“[T]he concept of a ‘chilling effect’ logically embraces every situation in which people are deterred from engaging in conduct, especially constitutionally protected conduct, by fear of prosecution due to the costs or risks of defending a lawsuit.”); id. at 884 n.192 (noting that justification for First Amendment overbreadth doctrine “would support a doctrine of equal sweep in cases involving alleged infringements of other fundamental rights”); Schauer, supra note 36, at 690 (“[I]nvidious chilling of constitutionally protected activity . . . can occur not only when activity shielded by the [F]irst [A]mendment is implicated, but also when any behavior safeguarded by the Constitution is unduly discouraged.”).


88 See Dorf, supra note 13, at 276 (stressing that Casey “applied ‘substantial overbreadth’ analysis” in deciding to invalidate the husband notification provision); Isserles, supra note 25, at 458 (observing that “Casey employed some version of the overbreadth doctrine in facially invalidating the spousal notification provision”).

89 Casey, 505 U.S. at 893.

90 Id. at 895. Casey, however, was not the first case to invalidate a law because most of its applications were unconstitutional. See Hodgson v. Minnesota, 497 U.S. 417, 460 (1990) (O’Connor, J., concurring) (invalidating two-parent notification statute on its face because
resembles substantial overbreadth analysis. Both rules demand that a plaintiff show that the statute is invalid in many or most of the statute’s applications and thereby ensure that facial invalidation occurs only when its benefits – protecting rights-holders from a chilling effect and forcing legislatures to regulate with precision – outweigh the costs of forbidding the government from enforcing statutory applications that, in the abstract, would be valid.91

Nearly ten years after Casey, in Stenberg v. Carhart, the Court facially invalidated a Nebraska statute that banned so-called “partial-birth” abortions except when necessary to save the woman’s life.92 The Court gave two reasons for striking the ban on its face. First, the statute lacked an exception permitting a “partial-birth abortion” when necessary to safeguard the woman’s health.93 Second, the statute prohibited the dilation and evacuation (“D&E”) procedure, the most common method for abortions after the first trimester, thereby imposing a substantial obstacle on the right to choose abortion.94

of its “broad sweep and its failure to serve the purposes asserted by the State in too many cases”) (emphasis added); City of Akron v. Akron Ctr. for Reprod. Health, Inc., 462 U.S. 416, 437-39 (1983) (invalidating on its face an ordinance requiring that second and third trimester abortions be performed in a hospital, despite the fact that such a requirement might be a valid health regulation as applied to abortions in the late second trimester or third trimester); see also Adler, Rights Against Rules, supra note 12, at 156 (arguing that most abortion cases prior to Casey “involved facial invalidation of rules that . . . had some morally acceptable applications”).

91 Casey’s strategic challenge differs from overbreadth challenges in one important respect. No standing exception exists in privacy cases (as opposed to free speech cases); only a protected plaintiff can bring a facial challenge. See H.L. v. Matheson, 450 U.S. 398, 406 (1981) (holding that a minor who failed to prove that she was mature or emancipated, and therefore lacked standing, could not challenge parental notification statute). The absence of the standing exception, it would seem, reflects a lack of need; an automatic standing exception is unnecessary for two reasons. First, physicians can raise the rights of their patients, so an automatic standing exception is unnecessary to get around traditional standing law. See Akron Ctr. for Reprod. Health, Inc., 462 U.S. at 440 n.30 (1983) (noting that “physician plaintiff . . . has standing to raise the claims of his minor patients”). Second, physicians have an economic incentive to challenge restrictive abortion laws, making the fear that a challenge will not be forthcoming unrealistic. Indeed, virtually every abortion case decided by the Court since Roe v. Wade, 410 U.S. 113 (1973), has involved physicians or clinics as plaintiffs. Once a physician brings suit – as often occurs – Casey’s strategic rule takes over, requiring facial invalidation of a statute that imposes an undue burden on a large fraction of women.

93 Id. at 930-38 (holding that Nebraska’s statute violates the Constitution because it “lacks any exception ‘for the preservation of the . . . health of the mother’” (quoting Planned Parenthood of Se. Pa. v. Casey, 505 U.S. 833, 879 (1992))); id. at 947-48 (O’Connor, J., concurring) (“[T]he Nebraska statute is inconsistent with Casey because it lacks an exception for those instances when the banned procedure is necessary to preserve the health of the mother.”).
94 Id. at 930, 938-44 (holding that Nebraska’s statute violates the Constitution because “it
The first ground for striking the Nebraska statute – the lack of a health exception – is a categorical defect of the sort *Salerno* seems to demand. The statute lacks a constitutionally required exception; therefore, it is unconstitutional on its face. But should that be enough to invalidate the entire statute? The lack of a health exception does not render the statute invalid as applied to all individuals; it only affects women for whom an abortion procedure defined under the law as a partial-birth abortion is necessary to protect their health. So why not enjoin the law only as applied to these women?95

The Court rejects this approach, most likely because the second ground for invalidating the law involves not merely the exception to the ban, but the ban itself. Echoing overbreadth reasoning, the Court invalidates the ban because it prohibits the procedure most commonly used after the first trimester (D&E), thereby chilling physicians from performing second trimester abortions and erecting a substantial obstacle to the abortion choice. By impermissibly banning the most common second trimester abortion procedure, the statute imposes an unconstitutional burden in a substantial number of cases, and thus cannot stand. It does not matter that the statute could be applied validly to the rarer dilation and extraction (“D&X”) procedure.96 In a *large majority* of cases, the statute is unconstitutional, and thus, as in *Casey*, the statute is void on its face.

What explains the *Casey* rule? The Court does not elaborate, but its reasons seem apparent. As in free speech cases, courts invalidate statutes on their face to protect absent third parties from the chilling effect that results from the threat of sanctions under the challenged statute. As in free speech cases, facial

95 Indeed, Justice Thomas faulted the Court for striking down the law for lack of a health exception when this failing only affected a small number of statutory applications. *Id.* at 1019-20 (Thomas, J., dissenting) (criticizing the Court for striking down Nebraska’s ban on “partial birth abortion” even though “there is no ‘large fraction’ of . . . women who would face a substantial obstacle to obtaining a safe abortion because of their inability to use this particular procedure”).

96 Apparently, a majority of the Court would permit a ban on the D&X procedure. The four dissenters, who voted to uphold the Nebraska ban, read the statute as limited to this procedure, *id.* at 972-79 (Kennedy, J., dissenting) (insisting that Nebraska’s statute, as written, “applies only to the D&X procedure”); *id.* at 989-97 (Thomas, J., dissenting) (arguing that Nebraska’s statute “does not prohibit the D&E procedure,” only the D&X procedure, and is therefore valid), and Justice O’Connor opined that a ban narrowly tailored to the D&X procedure would be constitutional. *Id.* at 950-51 (O’Connor, J., concurring) (“[A] ban on partial birth abortion that only proscribed the D&X method . . . would be constitutional in my view.”).
invalidation eliminates the chilling effect that would result if case-by-case adjudication were used to separate valid applications of the statute from unconstitutional applications. *Casey’s* rule is that state action deterring a woman’s exercise of her right of reproductive choice is unconstitutional, and the chilling effect strategy follows naturally from this definition of the right. *Casey’s* facial invalidation of the husband notice provision is a perfect example. As the Court pointed out, “the significant number of women who fear for their safety . . . are likely to be deterred from procuring an abortion as surely as if the Commonwealth had outlawed abortion in all cases.” To prevent this chilling effect, the law must be invalidated on its face. Case-by-case remedies, as in free-speech cases, will be ineffectual because they fail to protect third parties from the law’s chilling effect.

Indeed, the reasons to distrust case-by-case adjudication are even stronger in abortion cases than in many First Amendment overbreadth cases. Given the time pressures of terminating an unwanted pregnancy, it is unreasonable to expect plaintiffs – whether physicians or women seeking abortions – to bring as-applied claims; there is simply too little time to litigate them. Given this time constraint, if case-by-case litigation is required, women will inevitably be forced into motherhood. To prevent this chilling effect, the Court invalidates on its face any abortion statute that poses an undue burden in a large fraction of cases. As in First Amendment overbreadth doctrine, constitutional violations in a substantial number of cases plus a defect in case-by-case adjudication – the chilling effect – lead to facial invalidation.

### B. *Fundamental Rights Vagueness Doctrine*

Unlike First Amendment overbreadth or privacy doctrine, the Court’s vagueness jurisprudence does not guarantee substantive rights; instead, it

---

97 Planned Parenthood of Se. Pa. v. *Casey*, 505 U.S. 833, 894 (1992); *see also id.* at 897 (“Whether the prospect of notification itself deters such women from seeking abortions, or whether the husband . . . prevents his wife from obtaining an abortion until it is too late, the notice requirement will often be tantamount to the veto held unconstitutional in *Danforth*.”).

98 *Stenberg*, indeed, explicitly embraced the chilling effect rationale. The partial-birth abortion ban, the Court reasoned, was facially invalid because its sweeping definition would chill doctors from performing abortions in too many cases:

[U]sing this law some present prosecutors and future Attorneys General may choose to pursue physicians who use D&E procedures, the most commonly used method for performing previability second trimester abortions. All those who perform abortion procedures using that method must fear prosecution, conviction, and imprisonment. The result is an undue burden upon a woman’s right to make an abortion decision.

530 U.S. at 945-46.

99 *See Case*, 505 U.S. at 942 (Blackmun, J., concurring in part and dissenting in part) (criticizing the dissent for failing to explain “how a battered woman is supposed to pursue an as-applied challenge”); Dorf, *supra* note 13, at 270 (“Due to pregnancy’s temporary nature, pregnant women may find case-by-case legal relief from abortion restrictions particularly impractical.”).
shields individuals from state laws that are framed in unclear and ambiguous terms.\textsuperscript{100} Despite this difference, vagueness doctrine, like First Amendment and privacy law, has long been concerned that legislation might chill the exercise of substantive constitutional rights.\textsuperscript{101} Because individuals will often respond to a law’s vague prohibitions by “‘steer[ing] far wider of the unlawful zone’ than if the boundaries of the forbidden areas were clearly marked,”\textsuperscript{102} a vague law that implicates constitutional liberties will deter people from exercising their rights.\textsuperscript{103} Because people are reluctant to face the prospect of criminal and civil liability, “[t]he threat of sanctions may deter . . . almost as potently as the actual application of sanctions.”\textsuperscript{104} To prevent this chilling effect, the Court will facially invalidate a vague law, even when the law is clear in some of its applications.\textsuperscript{105} And it will do so across the spectrum of

\textsuperscript{100} See Grayned v. City of Rockford, 408 U.S. 104, 108 (1972) (“It is a basic principle of due process that an enactment is void for vagueness if its prohibitions are not clearly defined.”).

\textsuperscript{101} See Kolender v. Lawson, 461 U.S. 352, 358 n.8 (1983) (noting that vagueness and overbreadth are “logically related and similar doctrines”).

\textsuperscript{102} Grayned, 408 U.S. at 109 (quoting Baggett v. Bullitt, 377 U.S. 360, 372 (1964)) (citation omitted).

\textsuperscript{103} See Fallon, supra note 36, at 904 (observing that vagueness doctrine seeks to prevent “chilling of constitutionally protected expression”); Note, The Void-For-Vagueness Doctrine in the Supreme Court, 109 U. PA. L. REV. 67, 75 (1960) (arguing that vagueness doctrine creates “an insulating buffer zone of added protection at the peripheries of several of the Bill of Rights freedoms”); id. at 80 (warning that a vague law poses the “danger that the state will get away with more inhibitory regulation than it has a constitutional right to impose, because persons at the fringes of amenability to regulation will rather obey than run the risk of erroneous constitutional judgment”).


\textsuperscript{105} As in privacy cases, vagueness facial challenges are always first-party claims; no standing exception exists. Vill. of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 495 (1982) (“A plaintiff who engages in some conduct that is clearly proscribed cannot complain of the vagueness of the law as applied to the conduct of others.”). What justifies this distinction between vagueness and overbreadth cases? After all, the same necessity that justifies the standing exception in speech cases would seem to apply to a person subject to a vague rights-restricting statute. Laurence H. Tribe, American Constitutional Law § 12-32, at 1035 (2d ed. 1988) (“Those whose expression is ‘chilled’ by the existence of an overbroad law or unduly vague statute cannot be expected to adjudicate their own rights, lacking by definition the willingness to disobey the law.”). Perhaps the reason is that the due process concerns at the heart of vagueness doctrine are conceptualized as personal rights that can only be vindicated by one actually uncertain about the law’s application to him or her, whereas First Amendment rights are held collectively by all speakers, so that any speaker may raise an overbreadth claim. But this distinction, even if it makes sense of the Court’s cases, is deeply unsatisfying. The whole point of providing a standing exception is strategic; it is a way to avert a chilling effect. From the standpoint of personal rights, neither the unprotected speaker nor the person on notice should have a claim
constitutional rights.  

Keyishian v. Board of Regents exemplifies the chilling effect strategy in operation. There, the Court struck down as overly vague New York laws that provided for the removal of state teachers who uttered or committed any seditious or treasonable words or acts. “The crucial consideration is that no teacher can know just where the line is drawn between ‘seditious’ and non-seditious utterances and acts.” In invalidating the statute on its face, the Court emphasized the chilling effect such vague terms would cause, describing the scheme as “a highly efficient in terrorem mechanism.” As the Court noted,

It would be a bold teacher who would not stay as far as possible from utterances or acts that might jeopardize his living. . . . The danger of that chilling effect upon the exercise of vital First Amendment rights must be guarded against by sensitive tools which clearly inform teachers what is being proscribed. Preventing a chilling effect drove the Court to invalidate the measure on its face. Certainly, no teacher would express surprise at being removed for exhorting students to engage in violent acts against the government. The problem was that the statute provided no clear dividing line between seditious and permissible speech or action, and reliance on case-by-case adjudication to flesh out the statute’s meaning over time would inevitably lead to self-censorship and suppression of speech in the interim. Because of this defect in case-by-case adjudication, the Court rightly turned to facial invalidation. Facial invalidation prevented the chilling effect and forced the state to clarify its terms before it could apply the law to anyone.

to raise the rights of parties not before the Court. A standing exception exists in First Amendment overbreadth cases as a matter of strategy; if a standing exception is denied in vagueness cases, the denial should be based on some strategic reason. It is hard to see any such reason.

The Court has struck down vague abortion statutes to prevent them from chilling a woman’s right to choose. See Colautti v. Franklin, 439 U.S. 379, 394 (1979) (invalidating abortion statute because it “conditions potential criminal liability on confusing and ambiguous criteria,” and “therefore presents serious problems of notice, discriminatory application, and chilling effect on the exercise of constitutional rights”). And, as Tony Amsterdam – who is widely recognized as the author of The Void-For-Vagueness Doctrine in the Supreme Court, supra note 103 – has pointed out, the Court’s facial invalidation of economic regulations in the heyday of the Lochner era was designed to prevent vague laws from chilling the property rights Lochner protected. See Note, supra note 103, at 74 n.38 (collecting cases in which the Court applied vagueness doctrine to strike down economic regulations); id. at 77-80, 84-85.


Id. at 597.

Id. at 599.

Id. at 601.

Id.
A second strategy, rooted in First Amendment licensing cases, is also at work in vagueness cases. A vague law allows the government to take advantage of the law’s vague terms to discriminate in a manner not easily detected.\textsuperscript{112} To overcome this detection problem, the Court invalidates vague laws that invite such discrimination. In \textit{Gentile v. State Bar of Nevada}, for example, the Court invalidated on vagueness grounds a rule that gave authorities too much discretion to discipline lawyers based on the content of their speech.\textsuperscript{113} The Court emphasized that “[t]he question is not whether discriminatory enforcement occurred here, and we assume it did not, but whether the Rule is so imprecise that discriminatory enforcement is a real possibility.”\textsuperscript{114} Given the difficulty of proving actual discriminatory enforcement, the Court used facial invalidation to protect speech and equality, rather than call for a case-by-case examination of whether discrimination occurred. As in overbreadth licensing cases, facial invalidation is appropriate because a case-by-case approach that demanded a showing of discriminatory enforcement would fail to detect many unconstitutional acts.\textsuperscript{115}

Of course, to hold a statute void for vagueness in fundamental rights cases – whether because of a chilling effect or the risk of discriminatory enforcement – it is not enough to show a lack of clarity around the margins of a statute. As in overbreadth and privacy doctrine, strategic facial invalidation in vagueness cases involves a substantiality component. Unless the vagueness runs to a large majority of a statute’s applications, a facial challenge will not lie.\textsuperscript{116} \textit{Kolender v. Lawson}\textsuperscript{117} represents one formulation of a substantiality requirement. Noting that “we have traditionally viewed vagueness and overbreadth as logically related and similar doctrines,” the Court explained that vagueness doctrine permits a “facial challenge if a law reaches ‘a substantial amount of constitutionally protected conduct.’”\textsuperscript{118} But this seems to ask the

\begin{enumerate}
  \item \textsuperscript{112} See \textit{Gentile v. State Bar of Nev.}, 501 U.S. 1030, 1051 (1991) (“The prohibition against vague regulations of speech is based in part on the need to eliminate the impermissible risk of discriminatory enforcement.”).
  \item \textsuperscript{113} Id.
  \item \textsuperscript{114} Id.; see also id. at 1082 (O’Connor, J., concurring) (“[A] vague law offends the Constitution because it . . . creates the possibility of discriminatory enforcement.”).
  \item \textsuperscript{115} See supra notes 57-59 and accompanying text (discussing First Amendment licensing cases in which the Court rejected the case-by-case approach for fear that discrimination would “pass under the judicial radar”).
  \item \textsuperscript{116} See \textit{Hill v. Colorado}, 530 U.S. 703, 733 (2000) (“[S]peculation about possible vagueness in hypothetical situations . . . will not support a facial attack on a statute when it is surely valid ‘in the vast majority of its intended applications.’” (quoting United States v. Raines, 362 U.S. 17, 23 (1960))); Cal. Teachers Ass’n v. State Bd. of Educ., 263 F.3d 888, 897 (9th Cir. 2001) (“[U]ncertainty at a statute’s margins will not warrant facial invalidation . . . .”).
  \item \textsuperscript{117} 461 U.S. 352 (1983).
  \item \textsuperscript{118} Id. at 358 n.8 (quoting \textit{Vill. of Hoffman Estates v. Flipside, Hoffman Estates, Inc.}, 455 U.S. 489, 494 (1982)); see also \textit{City of Houston, Tex. v. Hill}, 482 U.S. 451, 459 (1987)
\end{enumerate}
wrong question. In a vagueness facial challenge, the question should not be how much protected conduct the statute covers; that question pertains to the different claim that the statute violates substantive constitutional rights. Rather, the question should be how far the statutory vagueness extends. Does the vagueness infect all applications, in which case the statute should fail under any standard? Does the vagueness infect a substantial number of cases, such that facial invalidation is necessary to prevent discrimination or a chilling of constitutional rights? Or does the vagueness only affect isolated applications of the law, in which case a facial challenge should fail, forcing parties to resort to as-applied remedies?

More recently, a plurality of the Court called for facial invalidation of criminal laws “[w]hen vagueness permeates the text” of the statute. The idea behind the “permeates the text” rule is that where a statute’s vagueness is pervasive, and not merely limited to individual outlying cases, courts should entertain facial challenges. Akin to the demand for substantial overbreadth, or Casey’s “large fraction” of applications test, the requirement that vagueness permeate the text ensures that the unconstitutional vagueness infects a substantial number of the statute’s applications. Where vagueness permeates the text, rejecting a facial challenge and forcing parties to resort to case-by-case adjudication risks an impermissible chilling effect on constitutional rights. The requirement of permeation ensures that the chilling effect will be widespread, not simply isolated cases that could be dealt with on a case-by-case basis.

Like the tests used in overbreadth and privacy cases, the (explaining that vague criminal statutes “that make unlawful a substantial amount of constitutionally protected conduct may be held facially invalid even if they also have legitimate application”).

City of Chicago v. Morales, 527 U.S. 41, 55 (1999) (“[Chicago’s Gang Congregation Ordinance] is a criminal law that contains no mens rea requirement and infringes on constitutionally protected rights. When vagueness permeates the text of such a law, it is subject to facial attack.”).

The Morales plurality qualified its test in two ways. It subjected the criminal statute to “facial attack” because “it contain[ed] no mens rea requirement” and “infringe[ed] on constitutional rights.” Id. The mens rea requirement is closely related to permeation. Where a criminal statute contains a mens rea requirement, it is unlikely that the statutory vagueness will be pervasive. As the Court observed in Screws v. United States:

[T]he requirement of a specific intent to do a prohibited act may avoid those consequences to the accused which may otherwise render a vague or indefinite statute invalid. . . . The requirement that the act must be willful or purposeful may not render certain . . . a statutory definition of the crime which is in some respects uncertain. But it does relieve the statute of the objection that it punishes without warning an offense of which the accused was unaware.” 325 U.S. 91, 101-02 (1945). But, as Screws suggests, even with a mens rea requirement, vagueness may still permeate a statute where the definition of the crime imposes no ascertainable standard of guilt. Thus, the touchstone should simply be permeation; there should be no separate requirement relating to mens rea. For discussion of the second qualification, the demand that the statute infringe constitutional rights, see infra text
permeation requirement ensures that facial invalidation occurs only when its benefits are high and its costs low.

C. The Right to Vote

Strategic facial invalidations are also a tool used to protect voting rights. *Louisiana v. United States* is illustrative. There, the Court invalidated on its face a state law that required, as a pre-requisite for voter registration, that a person satisfy the registrars that he or she could understand and interpret a provision of the State or Federal Constitution. The law was invalid for two reasons, corresponding to the two strategies present in both free speech licensing and vagueness cases. First, the sweeping discretion reposed in registrars – what the Court called “not a test but a trap, sufficient to stop even the most brilliant man on his way to the voting booth” – would chill persons, particularly African-Americans, from seeking to register to vote. Second, the grant of excessive discretion invited discrimination that was difficult to police in individual cases; the test’s vagueness and subjectivity “practically place[ ] [the registrar’s] decision beyond the pale of judicial review; and he can enfranchise or disenfranchise voters at his own sweet will and pleasure without let or hindrance.”

Compare *Louisiana* with the Court’s earlier ruling in *Lassiter v. Northampton County Board of Elections*. In *Lassiter*, the Court upheld a requirement that an individual be able to read and write any section of the U.S. Constitution, as a condition of voter registration, explaining that “[t]he ability to read and write . . . has some relation to standards designed to promote

accompanying note 168 (arguing that a vague statute that gives the police excessive discretion to define criminal conduct should be facially invalidated even if it does not infringe or chill substantive constitutional rights).


122 *Id.* at 153 (“[S]tatutes which require voters to satisfy registrars of their ability to ‘understand and give a reasonable interpretation of any section’ of the Federal or Louisiana Constitution violate the Constitution.”).

123 In fact, the Court specifically referenced vagueness doctrine in invalidating the measure. *Id.* at 153 (“Many of our cases have pointed out the invalidity of laws so completely devoid of standards and restraints.”).

124 *Id.*

125 See *id.* at 150 (“[T]he existence of the test as a hurdle to voter qualification has in itself deterred and will continue to deter Negroes from attempting to register in Louisiana.”).

126 Of course, the vagueness of the standard does not make discrimination impossible to police. See *id.* at 151 (finding overwhelming evidence of discriminatory application). But this only shows that heavy-handed discrimination will not always be able to hide behind a vague rule. Where subtler means of discrimination are employed, a sweeping grant of discretion can nevertheless make discrimination exceedingly difficult to detect.

127 *Id.* at 152 (citation omitted).

intelligent use of the ballot.”

Lassiter, in my view, shows that some applications of Louisiana’s interpretation test would be perfectly valid. Take, for example, an applicant who responds “I don’t understand that” when questioned by registrars about the First Amendment. A fair reading of Lassiter is that no as-applied constitutional claim would lie on those facts. Against the backdrop of Lassiter, Louisiana’s strategic move becomes clear. Even though the state might permissibly apply the interpretation test to deny a voter’s application for registration for legitimate, non-discriminatory reasons, the measure must be invalidated on its face because leaving the test in place would chill a substantial number of persons from seeking to register to vote, and would result in discrimination likely passing undetected.

What about unconstitutional purpose? The provision in Louisiana was, quite obviously, designed and enforced with the intention of disenfranchising African-American voters. Is illicit purpose what drove the facial invalidation, not strategic concern about a chilling effect or the difficulty of proving discrimination on a case-by-case basis? The purpose reading is difficult to square with what the Court actually said in Louisiana. The Court’s rationale rested on the character of the legislative action. The test was invalid on its face because of the discretion reposed in state officials to deny persons the right to vote. That character rendered the statute invalid on its face, whether or not the discretion was given for a proper purpose. Had the discretion not been given for racial reasons, but to “promote intelligent use of the ballot,” facial invalidation would still follow. “The cherished right of people . . . to vote cannot be obliterated by the use of laws like this, which leave the voting fate of a citizen to the passing whim or impulse of an individual registrar.” Such discretion is incompatible with the right to vote, not because the discretion is bestowed for an invalid purpose, but because of the dangers of discretion: it chills the exercise of rights and makes discrimination hard to uncover on a case-by-case basis. These strategies justify facial invalidation.

129 Id. at 51.

130 See United States v. Louisiana, 225 F. Supp. 353, 363-81 (E.D. La. 1963) (three-judge panel) (reviewing Louisiana voting laws since 1724 to show that the interpretation test is “rooted in the State’s historic policy and the dominant white citizens’ firm determination to maintain white supremacy in state and local government by denying to Negroes the right to vote”).

131 In fact, when the Court refused to facially invalidate a scheme giving jury commissioners excessive discretion in selecting jurors, it distinguished Louisiana based on the unconstitutional purpose of the interpretation test. See Carter v. Jury Comm’n, 396 U.S. 320, 336 & n.41 (1970).

132 Lassiter, 360 U.S. at 51.

D. The Right to Travel

Right to travel cases provide a final example of the chilling effect strategy. The seminal case in this line is *Aptheker v. Secretary of State*, in which the Court invalidated a statute prohibiting any member of a Communist organization from applying for or using a passport.\textsuperscript{134} *Aptheker* sounds very much like a First Amendment case because the statute conditioned the right to obtain a passport on surrendering the First Amendment right to join a political party. But the Court did not hold the statute invalid under the First Amendment; it found that the statute unconstitutional because it prohibited too much travel.\textsuperscript{135} The statute “swe[pt] too widely and too indiscriminately” and was “supported [only by] a tenuous relationship between the bare fact of organizational membership and the activity Congress sought to proscribe.”\textsuperscript{136} In sum, the statute was substantially overbroad, and thus facially invalid.

The Court declined the dissent’s invitation to uphold the statute as applied to the plaintiffs, both high-ranking party leaders, and to put off the question of whether the statute might be invalid as applied to others.\textsuperscript{137} The Court gave two reasons for judging the law on its face. First, any possible saving construction would involve substantial rewriting of the law and would only trade the statute’s overbreadth for vagueness.\textsuperscript{138} Second, the Court specifically invoked First Amendment overbreadth doctrine, holding that challenges to statutes curtailing constitutionally protected liberty should not be adjudicated on a case-by-case basis.\textsuperscript{139}

The Court’s first reason hardly answers Justice Clark’s dissent. Justice Clark, after all, did not propose a limiting construction. Rather, he would have rejected plaintiffs’ claim as applied, which would have left the statute in place and permitted future challenges by differently-situated claimants.\textsuperscript{140} The Court’s second rationale, its chilling effect theory, explains why the Court

\textsuperscript{134} 378 U.S. 500, 505 (1964) (finding § 6 of the Subversive Activities Control Act unconstitutional).
\textsuperscript{135} *Id.* at 517.
\textsuperscript{136} *Id.* at 514.
\textsuperscript{137} The dissent found irrelevant the sweeping nature of the restrictions – their application to unknowing members and to all travel – because plaintiffs were top party leaders. *Id.* at 523-25 (Clark, J., dissenting) (arguing that the statute should be held constitutional “[a]s applied to the prosecution of the Communist Party’s top dignitaries”). As Justice Clark argued:

[W]e have no ‘unknowing members’ before us. . . . All we have are irrational imaginings: a member of the Party might wish ‘to visit a relative in Ireland, or to read rare manuscripts in the Bodleian Library of Oxford University . . . .’ But no such party is here and no such claim is asserted. It will be soon enough to test this situation when it comes here.

*Id.* at 525 (Clark, J., dissenting).
\textsuperscript{138} *Id.* at 516.
\textsuperscript{139} *Id.* at 516-17.
\textsuperscript{140} See *id.* at 523-25 (Clark, J., dissenting).
rejects Justice Clark’s approach. To uphold the statute as applied to the party leaders, and leave the statute in place until a better claimant brought suit, would chill the constitutional rights of absent party members to travel. The statute would chill all manner of travel raising no legitimate security concerns, such as a trip “to visit a relative in Ireland, or to read rare manuscripts in the Bodleian Library of Oxford University,” as well as all travel by unknowing or uninvolved Communist Party members whose relationship to the party was tenuous at best. To prevent this chilling effect on travel, the Court allowed the top-ranking members to raise the constitutional rights of all affected party members, and invalidated the statute on its face as overbroad, thus preempting case-by-case adjudication. The chilling effect strategy is plainly at work here.

IV. THE EXCESSIVE DISCRETION STRATEGY

The second invalidation strategy – the excessive discretion strategy – is concerned with the harms caused by statutes that confer too much discretion on actors to violate constitutional commands. It figures in two quite distinct areas of law, vagueness and Establishment Clause cases, and, as we shall see, operates differently in each context. In vagueness cases, facial invalidation is a means to better protect against discriminatory enforcement of criminal laws; in Establishment Clause cases, facial invalidation aims to protect persons from impermissible religious inculcation before it occurs.

A. Excessive Discretion, Discrimination, and Vagueness

In the fundamental rights vagueness cases surveyed above, we saw two reasons to prefer facial invalidation to case-by-case adjudication: (i) to prevent a chilling effect on constitutional rights and (ii) to prevent state officials from

---

141 Id. at 512.
142 Although Aptheker appears to follow overbreadth doctrine, Marc Isserles argues that the case did not rely on that doctrine at all. See Isserles, supra note 25, at 413. Instead, in his view, the Aptheker Court simply applied strict scrutiny and found that the passport regulation was not narrowly tailored “to serve the State’s admittedly compelling interest of protecting national security.” Id. In other words, the Court struck down the statute because it was an invalid rule of law; the Court simply “followed the mode of analysis employed in . . . Salerno.” Id. Isserles, however, is too limited in his reading of overbreadth doctrine. Overbreadth, like constitutional law more generally, is concerned with the validity of rules, not merely their application, and the means by which the government justifies its incursions on protected activity. Indeed, virtually every Supreme Court overbreadth case which the challenger prevailed relied on the government’s failure to regulate with necessary precision. See, e.g., Ashcroft v. Free Speech Coal., 535 U.S. 234, 252-53, 258 (2002) (“The First Amendment requires a more precise restriction. For this reason, [the statute] is substantially overbroad and in violation of the First Amendment.”). A statute that is substantially overbroad, almost by definition, is one which lacks the precision that narrow tailoring requires. Substantial overbreadth and a lack of narrow tailoring thus go hand in hand. Isserles slights this key connection.
using broad grants of discretion to engage in discrimination, which might otherwise escape detection. Facial invalidation, however, should not depend on a showing that the law’s vagueness chills constitutional rights. The fact that a criminal law gives the police excessive discretion to define the law’s parameters justifies a strategic facial challenge.

*Kolender v. Lawson* and *City of Chicago v. Morales* guide our way. Properly read, these cases show that a facial challenge is appropriate where a criminal law gives police excessive discretion to define criminal conduct because excessive discretion poses a risk of discriminatory enforcement that cannot be effectively prevented on a case-by-case basis. Where vagueness “permeates the text” of the discretion-granting statute, and thus constitutional violations are pervasive, a facial challenge is a better means of implementing the Constitution because of the limits of case-by-case adjudication.

In *Kolender*, the Court invalidated on its face a criminal statute requiring persons validly stopped by the police to provide a “credible and reliable” identification. Explaining that the statute threatened to “arbitrarily suppress[] First Amendment liberties,” the Court invalidated the statute on its face, finding that the statute “vests virtually complete discretion in the hands of the police to determine whether the suspect has satisfied the statute and must be permitted to go on his way in the absence of probable cause for arrest.”

The Court specifically rejected the dissent’s call for a *Salerno*-like facial challenge rule, observing the close connection between vagueness and overbreadth, and dismissed the dissent’s assertion that the law was facially valid because it was unambiguous in cases in which the suspect refused to provide any identification. The statute’s constitutional flaw – its grant of

---

143 See *supra* Part III.B.
147 *Morales*, 527 U.S. at 55. For discussion of the “permeates the text” rule, see *supra* text accompanying notes 119-120.
148 *Kolender*, 461 U.S. at 352.
149 Id. at 358 (quoting Shuttlesworth v. City of Birmingham, 382 U.S. 87, 91 (1965)).
150 Id. at 358 n.8 (“[W]e have traditionally viewed vagueness and overbreadth as logically related and similar doctrines.”).
151 Id. at 372 (White, J., dissenting) (“Suppose . . . an officer requests identification information . . . during a valid *Terry* stop and the suspect answers: ‘Who I am is just none of your business.’ Surely the suspect would know from the statute that a refusal to provide any
near-total discretion to officers to determine whether the identification offered by a stopped suspect was “credible and reliable”\textsuperscript{152} – infected virtually all applications of the law and necessitated the statute’s facial invalidation.\textsuperscript{153}

\textit{Morales} follows closely on the heels of \textit{Kolender}. \textit{Morales} invalidated a Chicago ordinance that made it illegal for “criminal street gang members” to “loiter,” defined as “remaining in any one place with no apparent purpose,” with one another or other persons in a public place after being ordered by a police officer to disperse.\textsuperscript{154} In striking down the law on its face, the plurality explained that “the vagueness of this enactment makes a facial challenge appropriate. . . . It is a criminal law that contains no \textit{mens rea} requirement, and infringes on constitutionally protected rights. When vagueness permeates the text of such a law, it is subject to facial attack.”\textsuperscript{155} The plurality specifically rejected use of \textit{Salerno} in favor of its “permeates the text” rule.\textsuperscript{156} Under this rule, the plurality found the ordinance void on its face, observing that the ordinance’s definition of loitering failed to provide persons with adequate notice and gave excessive discretion to the police.\textsuperscript{157}

\textsuperscript{152} \textit{Id}. at 360. The Court explained:

It is clear that the full discretion accorded to the police to determine whether the suspect has provided a credible and reliable identification necessarily entrusts lawmaking to the moment-to-moment judgment of the policeman on the beat [and] furnishes a convenient tool for harsh and discriminatory enforcement by local prosecuting officials, against particular groups deemed to merit their displeasure, and confers on police a virtually unrestrained power to arrest and charge persons with a violation.\textit{Id}. (internal citations and quotation marks omitted).

\textsuperscript{153} \textit{Id}. at 361 (concluding that the statute was “unconstitutionally vague on its face because it encourages arbitrary enforcement by failing to describe with sufficient particularity what a suspect must do in order to satisfy the statute”).


\textsuperscript{155} \textit{Id}. at 55.

\textsuperscript{156} \textit{Id}. at 55 n.22 (“To the extent we have consistently articulated a clear standard for facial challenges, it is not the \textit{Salerno} formulation . . . . Since we . . . conclude that vagueness permeates the ordinance, a facial challenge is appropriate.”). The \textit{Morales} plurality could not garner five votes either to limit or overrule \textit{Salerno} or to adopt the “permeates the text” rule. Six Justices (the plurality and Justices O’Connor, Breyer, and Kennedy) did, however, agree that the loitering ordinance was void on its face. Unlike the plurality, Justice Breyer sought to shoehorn the case into \textit{Salerno}:

The ordinance is unconstitutional, not because a policeman applied this discretion wisely or poorly in a particular case, but rather because the policeman enjoys too much discretion in every case. And if every application of the ordinance represents an exercise of unlimited discretion, then the ordinance is invalid in all its applications.\textit{Id}. at 71 (Breyer, J., concurring). Neither Justice O’Connor nor Justice Kennedy addressed the appropriate facial challenge rule.

\textsuperscript{157} \textit{Id}. at 56-64. The majority of the Court joined the portion of Justice Stevens’ opinion concerning excessive discretion.
Both *Kolender* and *Morales* emphasize that constitutional freedoms are at risk because of the law’s vagueness, implicitly drawing on the chilling effect narrative to justify invalidating the laws on their face. But chilling effects cannot justify the Court’s invalidations. A more persuasive explanation lies in a second theory, rooted in the evils of excessive discretion.

The chilling effect argument proceeds as follows. The laws in *Kolender* and *Morales* denied individuals the free use of public streets by granting the police too much discretion to arrest and charge individuals with crimes. Rather than subjecting themselves to the risk of arrest and prosecution, individuals would steer far clear of the law’s boundaries. Thus, the law would chill their use of public streets and their freedom of movement. Facial invalidation was necessary to stop this chilling effect.

This argument is unpersuasive. It is easy to understand the chilling effect of a vague (or overbroad) law that targets a particular kind of conduct. Thus, in *Keyishian*, the Court explained that New York’s vague prohibition of “seditious” utterances would deter teachers from engaging in any kind of speech critical of the government. The vagueness in the ordinance at issue in

---

158 The loss of constitutional freedoms, however, does little actual work in the opinions. In *Kolender*, the Court mentions its concern that constitutional freedoms will be lost. But it says no more than that. *Kolender*, 461 U.S. at 358. In *Morales*, the plurality appears to give substantive constitutional rights a greater role in its analysis; it declares the Chicago ordinance “subject to facial attack,” because it “infringes on constitutionally protected rights.” *Morales*, 527 U.S. at 55. But this language is misleading. All the plurality says about substantive constitutional rights is that innocent loitering “is part of the ‘liberty’ protected by the Due Process Clause of the Fourteenth Amendment.” *Id.* at 53. It makes no finding that the ordinance should be invalid for violating this liberty, and, in fact, chides the dissenters for “incorrectly assum[ing] . . . that identification of an obvious liberty interest that is impacted by a statute is equivalent to finding a violation of substantive due process.” *Id.* at 53 n.19. So, as in *Kolender*, all that the Court says about substantive constitutional rights is that those rights are implicated by the statute’s vagueness.

159 There is another possibility to consider. Perhaps the Court is not relying on the chilling effect strategy at all, but is simply relying on the fact that vagueness scrutiny is stricter when constitutional rights are involved. See *Kolender*, 461 U.S. at 358 n.8 (“[W]e permit a facial challenge if a law reaches a ‘substantial amount of constitutionally protected conduct.’” (quoting Vill. of Hoffman Estates v. Flipside, Hoffman Estates, Inc., 455 U.S. 489, 494 (1982))); Vill. of Hoffman Estates, 455 U.S. at 499 (“[P]erhaps the most important factor affecting the clarity that the Constitution demands of a law is whether it threatens to inhibit the exercise of constitutionally protected rights.”)). But stricter scrutiny of a law’s terms does not necessarily translate into a demand for facial relief. That the standard is stricter simply means that a court will have less tolerance for unclear terms. Using the same strict standard, a court could choose to rely solely on case-by-case adjudication, or it could enforce its higher standard by granting facial relief against vague laws. Use of a stricter vagueness standard does not necessarily preclude insistence on case-by-case adjudication. *See Kolender*, 461 U.S. at 371-73 (White, J., dissenting) (agreeing that stricter vagueness standard was applicable but rejecting facial challenge because statute was not vague in all applications).
Morales, on the other hand, did not target any specific conduct at all, which greatly weakens any claim of chilling effect. Moreover, no criminality attached until the persons refused to disperse after being ordered to do so by the police. Any person concerned that his or her action was criminal could wait to see if the police ordered him or her to disperse. The chilling effect argument in Kolender fares no better. Because the law at issue required a valid Terry stop\textsuperscript{160} – meaning that the police needed an objectively reasonable suspicion of criminal conduct – before any request for identification could be made, the risk of chilling protected conduct was scant at best.

Instead, Kolender and Morales depend on a second strategy – one directed at the substantial risk of discrimination that excessive discretion creates, a chief concern of vagueness doctrine.\textsuperscript{161} A strong anti-discrimination principle, centrally concerned with discrimination against disfavored members of a community (particularly racial discrimination) has long animated vagueness doctrine.\textsuperscript{162}

Anti-discrimination norms explain why the Court in Kolender and Morales viewed facial invalidation as a better means of implementing the Constitution than case-by-case adjudication, and provide a more persuasive justification for the invalidations. A vague law that grants excessive discretion to police to define what constitutes criminal conduct creates a substantial risk of discrimination. It gives the police virtually unchecked discretion to decide who has violated community norms, and thus “furnishes a convenient tool for ‘harsh and discriminatory enforcement by local prosecuting officials, against particular groups deemed to merit their displeasure.’”\textsuperscript{163} Case-by-case adjudication cannot effectively prevent such discrimination because the vague terms of the law make discrimination very difficult to detect.\textsuperscript{164}

\textsuperscript{160} See Terry v. Ohio, 392 U.S. 1 (1968).

\textsuperscript{161} As Kolender explains:

Although [vagueness] doctrine focuses both on actual notice to citizens and arbitrary enforcement, . . . the more important aspect of vagueness doctrine “is not actual notice, but the other principal element of the doctrine – the requirement that a legislature establish minimal guidelines to govern law enforcement.” Where the legislature fails to provide such minimal guidelines, a criminal statute may permit “a standardless sweep [that] allows policemen, prosecutors, and juries to pursue their personal predilections.”


\textsuperscript{164} See Cole, supra note 163, at 1084 (“[V]ague laws invite selectivity and make discriminatory enforcement far more difficult to monitor.”); Jeffries, supra note 162, at 213 (“Greater conformity to the rule of law discourages resort to illegitimate criteria of selection
by-case adjudication, some unconstitutional discrimination will likely pass under the judicial radar.

To see why this is so, compare the statute invalidated in Kolender, which required a “credible and reliable” identification, with a statute that simply required a person to provide a government-issued identification or his or her name. In the latter case, the statute’s clear and objective terms provide a baseline against which to assess a claim of discrimination. We know exactly what information an officer can demand of a suspect he or she has stopped pursuant to Terry, and we are thus in a position to assess on a case-by-case basis the legality of any arrest. But in the former case, absent a limiting construction from the courts, we have no idea what a “credible and reliable” identification entails; the baseline – what “credible and reliable” means – is constantly shifting. Every police officer could have a different view of what sorts of identifications meet the test. One officer might require a passport or driver’s license; another might accept any government-issued identification but closely scrutinize the photograph; a third might be satisfied by any form of photo identification. In these circumstances, proving discrimination becomes exceedingly difficult. After the fact, an officer could defend virtually any arrest against a discrimination claim as a reflection of his or her high standard for a “credible and reliable” identification. Within the confines of state criminal justice systems, in which opportunities for discovery are extremely limited to say the least, a suspect arrested for failure to provide a “credible” identification has virtually no way to amass the evidence needed to prove discrimination.

A strategic facial challenge is thus necessary because a vague criminal law that gives excessive discretion to police to define criminal conduct presents a substantial risk of discrimination, which courts cannot effectively monitor and check using as-applied adjudication. Where vagueness “permeates the and enhances our ability to discover and redress such abuses when they occur.”); Note, supra note 103, at 80 (“Prejudiced, discriminatory or overreaching exercises of state authority may remain concealed beneath findings of fact impossible for the Court to redetermine when such sweeping statutes have been applied to the complex, contested fact constellations of particular cases.”).


166 A similar concern with excessive discretion and the difficulties of proving actual discrimination underlies the Supreme Court’s recent decision in Nev. Dep’t of Human Res. v. Hibbs, 538 U.S. 721 (2003), which upheld the abrogation of Eleventh Amendment immunity under the Family and Medical Leave Act of 1993 (“FMLA”). Hibbs upheld the FMLA’s provisions granting workers the right to twelve weeks of unpaid medical leave per year as valid anti-discrimination legislation to enforce the Equal Protection Clause. The statutory entitlement to medical leave was necessary to correct unconstitutional gender discrimination, the Court reasoned, because leave policies gave substantial discretion to supervisors to provide leave in individual cases, and those discretionary decisions were often infected by the stereotypical reasoning of supervisors. Id. at 732, 736. Because the
text" of the law, and is not merely limited to isolated cases, a facial challenge is a better tool to prevent discrimination than as-applied challenges. Facial invalidation is appropriate because a case-by-case approach that demanded a showing of discriminatory enforcement would fail to detect many unconstitutional acts.

*Kolender* and *Morales*, unfortunately, tried to force these cases into the chilling effect mold by emphasizing the statute’s impact on constitutional rights. And *Morales* went so far as to build into its facial challenge rule the requirement that the statute infringe substantive constitutional rights. But in these cases, any chilling effect was at best weak and cannot justify the facial invalidations. Rather, *Kolender* and *Morales* represent a second type of strategy in vagueness cases, an equal protection theory of vagueness. Under this equal protection theory of vagueness, a statute that gives the police excessive discretion to determine what constitutes criminal conduct is facially invalid where vagueness permeates the statutory text (even if valid in certain cases) because it invites discrimination that courts cannot effectively monitor and stop through case-by-case adjudication.

A critic, however, might object that the Court should not care about a risk of discrimination. After all, throughout constitutional law, when a law enforcement officer or other state actor has wide discretion to act, the Court generally permits such discretionary action, subject only to the rule of *Washington v. Davis*, that purposeful discrimination offends the Constitution. In all contexts where a law enforcement officer may properly exercise discretion, the risk always exists that unconstitutional racial discrimination will go undetected, but notwithstanding that risk, the Court holds potential plaintiffs to the burden of proving intentional racial (or other

result was “subtle discrimination that may be difficult to detect on a case-by-case basis,” Congress could create a right to medical leave and deny employers discretion over leave decisions as a prophylactic measure to prevent such “difficult to detect” gender-based discrimination. *Id.* at 736.

---


168 See *id.* It follows from my analysis that the “permeates the text” rule from *Morales* should not require a showing that a statute implicates, let alone infringes, substantive constitutional rights. Avoiding the risk of unchecked discrimination necessitates facial invalidation, whether or not a substantive constitutional right is targeted.


illicit) discrimination. Why should vagueness doctrine be any different? Rather than facially invalidating state laws because of a risk of discrimination, the Court should permit case-by-case review, giving defendants the opportunity to prove actual discrimination.

The answer rests on the close connection between vagueness doctrine, particularly the requirement that the legislature define the conduct it chooses to make criminal, and the rule of law. The foundation of criminal law and procedure is a clearly defined law. It is such a law that enables courts to assess whether the police have probable cause or reasonable suspicion for searches and seizures. It is such a law that allows judges and juries to assess guilt or innocence. A vague criminal statute that gives excessive discretion to define criminal conduct, at least in the absence of a limiting construction, allows law enforcement officials to make the law as they enforce it; it is akin to a delegation of lawmaking power to the police, an invitation to "round up the usual suspects." Such a statute represents a breakdown of the rule of law; it not only carries with it the threat that police will use their power to discriminate against disfavored persons, it infects the criminal process system-wide. For this reason, judicial control, and thus facial invalidation, is of particular importance. These features, a threat to the rule of law and a fear of hard-to-detect discrimination, justify the Court’s embrace of strategic invalidation.

There is also a crucial difference in the Court’s valuation of discretion. The Court recognizes the importance of giving substantial discretion to the police

---

171 See, e.g., Whren, 517 U.S. 806 (placing burden on plaintiff to prove police inappropriately initiated traffic stop); Armstrong, 517 U.S. 456 (placing burden on plaintiff to prove prosecutor discriminated in bringing criminal charges); McCleskey, 481 U.S. 279 (placing burden on plaintiff to prove prosecutor discriminated in deciding to seek and impose death penalty); Batson, 476 U.S. 79 (placing burden on plaintiff to prove prosecutor discriminated while exercising peremptory challenges); Carter, 396 U.S. 320 (placing burden on plaintiff to prove jury commissioner discriminated in selecting persons for jury service).

172 The rule of law, John Jeffries explains, “means that the agencies of official coercion should, to the extent feasible, be guided by rules – that is, by openly acknowledged, relatively stable, and generally applicable statements of proscribed conduct.” Jeffries, supra note 162, at 212.

173 See generally Lanzetta v. New Jersey, 306 U.S. 451, 453 (1939) (“No one may be required at peril of life, liberty, or property to speculate as to the meaning of penal statutes.”); Connally v. Gen. Constr. Co., 269 U.S. 385, 391 (1926) (“A statute which either forbids or requires the doing of an act in terms so vague that men of common intelligence must necessarily guess at its meaning and differ as to its application, violates the first essential of due process of law.”).

174 See Jeffries, supra note 162, at 215 (“The power to define a vague law is effectively left to those who enforce it, and those who enforce the penal law characteristically operate in settings of secrecy and informality, often punctuated by a sense of emergency, and rarely constrained by self-conscious generalization of standards.”).
to search and seize suspects, to prosecutors to make decisions about whom to charge with crimes, and to attorneys to choose jurors.\footnote{See, e.g., \textit{Armstrong}, 517 U.S. at 464; \textit{McCleskey}, 481 U.S. at 296, 311; \textit{Batson}, 476 U.S. at 91, 98-99.} Not so with a statute that gives excessive discretion to police to define criminal conduct. The job of defining criminal conduct is a legislative function. A breakdown in the rule of law occurs because the legislature has abdicated its responsibility by entrusting lawmakers power to the police, even when the legislature could have drafted a clearer law.\footnote{Compare \textit{Kolender v. Lawson}, 461 U.S. 352, 361 (1983) (emphasizing that clearer terms were neither “impossible or impractical”) \textit{with United States v. Petrillo}, 332 U.S. 1, 7-8 (1947) (“Clearer and more precise language might have been framed by Congress to express what it meant by ‘number of employees needed.’ But none occurs to us, nor has any better language been suggested. . . . [T]he Constitution does not require impossible standards.”).} Excessive discretion is thus gratuitous discretion, and, for that reason, the cost-benefit calculus weights in favor of facial invalidation. The benefits of facial invalidation – protection against hard-to-detect discrimination – outweigh the costs of denying discretion to the police. Facial invalidation, thus, properly forces legislatures to draft criminal laws with greater precision.\footnote{These considerations, I think, explain the distinctiveness of vagueness doctrine. Of course, we might ask whether the balance the Court struck is the right one and whether excessive discretion should be grounds for strategic invalidation in other contexts as well. For example, consider \textit{McCleskey}. Justice Brennan’s dissent makes a powerful case that the risk of discrimination in death penalty cases demands invalidation of current death penalty schemes because too much discrimination will go unchecked since it is too difficult to detect on a case-by-case basis. \textit{McCleskey}, 481 U.S. at 322, 323 (Brennan, J., dissenting) (focusing on “the risk of the imposition of an arbitrary sentence, rather than the proven fact of one” because of “the difficulty of divining the jury’s motivation in an individual case”). Justice Brennan’s argument, in a sense, is that the Court should employ the same kind of invalidation strategy in Eighth Amendment cases that it does in vagueness doctrine.}

B. \textit{Excessive Discretion in Establishment Clause Challenges}

Establishment Clause facial challenges, at first glance, seem an odd candidate for the excessive discretion strategy. Unlike vagueness cases, in which the doctrine is full of talk of excessive discretion, Establishment Clause cases hardly, if ever, use the term. Nevertheless, a unifying theme links the two areas of constitutional law. In both, the Court holds facially invalid governmental grants of power (i.e., discretion to act) rather than review case-by-case whether the actor exercised the granted discretion in an unconstitutional manner. Thus, in vagueness cases, the fact that the government has granted police officers excessive discretion to define criminal conduct renders a statute facially void for vagueness, thereby preempting the question whether, in a particular case, the police officer abused his or her discretion. Likewise, in Establishment Clause cases, the Court holds facially
invalid statutes and programs that grant too much discretion to advance religion, rather than engaging in case-by-case review to see whether the actors used their authority to indoctrinate or otherwise impermissibly inculcate religion. 178

Facial invalidation in both settings is strategic; the Court facially invalidates the grant of discretion because case-by-case review is too weak to protect the underlying constitutional rights. But the reason that case-by-case review is insufficient differs sharply in vagueness and Establishment Clause cases. Whereas Kolender and Morales mandate facial review because of the risk that hard-to-detect discrimination will go unchecked under case-by-case review, the Establishment Clause cases mandate facial review because facial invalidation of the grant of discretion is the only way to prevent constitutional violations before they occur; as-applied review is inadequate because it comes too late, after impermissible religious inculcation has occurred.

Let us now turn to the cases. In Santa Fe Independent School District v. Doe, the Supreme Court facially invalidated a Texas school district’s policy of granting students the power to elect a speaker to give a message or statement before high school football games. 179 The school’s policy provided for a two-step electoral process: first, students would vote on whether to have a student deliver a solemnizing message or invocation; if the students chose to have a student speaker, a second election would be held to choose that speaker, who would then decide what message to deliver. 180

The Court held the policy invalid for two reasons. First, the policy’s purpose was to encourage religion; facial invalidation flowed inexorably from the Court’s finding of an unconstitutional purpose. 181 In this regard, Santa Fe is unexceptionable. Second, and importantly for our purposes, the policy gave students too much power to inculcate religion. The policy “allow[ed] the student majority to control whether students of minority views are subjected to a school-sponsored prayer.” 182 Although the Court never used the phrase “excessive discretion,” the policy’s constitutional flaw inhered in the sweeping power granted to students. A voting majority of students, under the policy, had the unfettered discretion to elect a religious speaker to give a prayer at school

178 In Establishment Clause facial challenges, the key question involves the effects inquiry. Where a statute has an unconstitutional purpose, the illicit purpose infects all applications of the law, making facial invalidation proper even under Salerno.  See Isserles, supra note 25, at 440-41 (arguing that courts that find the government’s purpose improper must necessarily invalidate a statute on its face); see also Dorf, supra note 13, at 279-81 (discussing the Court’s Establishment Clause jurisprudence that allows for facial invalidation of statutes with an unconstitutional purpose).


180 Id. at 296-97.

181 Id. at 316 (“[T]he simple enactment of this policy, with the purpose and perception of school endorsement of student prayer, was a constitutional violation. We need not wait for the inevitable to confirm and magnify the constitutional injury.”).

182 Id. at 316 n.23.
football games, leading to the kind of government-supported indoctrination which the Establishment Clause specifically prohibits.\textsuperscript{183} The grant of such limitless discretionary power to inculcate religion demanded facial invalidation.\textsuperscript{184}

Chief Justice Rehnquist, in dissent, took the majority to task for invalidating the policy on its face in defiance of \textit{Salerno}. \textit{Salerno} demanded a showing that the policy had no valid applications, and no overbreadth claim could be made in Establishment Clause cases. “No speech will be ‘chilled’ by the existence of a government policy that might unconstitutionally endorse religion over nonreligion. Therefore, the question is not whether the district’s policy may be applied in violation of the Establishment Clause, but whether it inevitably will be.”\textsuperscript{185} The dissenters found this high threshold unmet. While recognizing that the policy could “lead to a Christian prayer in 90 percent of the football games,”\textsuperscript{186} the dissent saw other possible ways in which the policy might play out. “[I]t is possible that the students might vote not to have a pregame speaker, in which case there would be no threat of a constitutional violation. It is also possible that the election would not focus on prayer, but on public speaking ability or social popularity.”\textsuperscript{187} Because of these varying possibilities, \textit{Salerno}, they argued, mandated rejection of the facial challenge.\textsuperscript{188}

But the dissenters’ error was in assuming that preventing the chill caused by an overbroad speech restriction was the only basis for departing from \textit{Salerno}. \textit{Santa Fe} exemplifies a different strategy: eliminating, before the fact, broad grants of discretion that make possible impermissible religious inculcation, rather than reviewing, after the fact and on a case-by-case basis, whether the discretion was used in an unconstitutional manner.

Giving students in the majority the unfettered power to foist their religion on the minority rendered the policy defective on its face, even if the election

\textsuperscript{183} \textit{Id.} at 316 (“Th[e] policy empowers the student body majority with the authority to subject students of minority views to constitutionally improper messages. The award of that power alone, regardless of the students’ ultimate use of it, is not acceptable.”).

\textsuperscript{184} See \textit{id.} (“This policy . . . does not survive a facial challenge because it impermissibly imposes upon the student body a majoritarian election on the issue of prayer.”).

\textsuperscript{185} \textit{Id.} at 318 (Rehnquist, C.J., dissenting). Chief Justice Rehnquist was certainly correct that the rationale for overbreadth – the chilling effect strategy – was inapplicable. Unlike the freedom of speech guarantee, the Establishment Clause does not protect an individual’s right to engage in any conduct at all. The Establishment Clause, instead, is a shield; it allows individuals to stop the government from favoring or endorsing religion. As an immunity against government establishments, not a conduct right, it is senseless to talk of a chilling effect on Establishment Clause rights. There is simply no conduct to be chilled.

\textsuperscript{186} \textit{Id.} at 321 (Rehnquist, C.J., dissenting).

\textsuperscript{187} \textit{Id.}

\textsuperscript{188} \textit{Id.} at 326 (“The policy at issue here may be applied in an unconstitutional manner, but it will be time enough to invalidate it if that is found to be the case.”).
process might be used to elect a student to speak on a secular topic. Absent a facial challenge to the policy, the Court implicitly concluded, substantial constitutional violations were likely to occur. Although the Court did not explicitly mention the substantiality requirement, it is critical to understanding why facial invalidation was proper. Were the risk that students would be subjected to religious exercises a lesser one, a case-by-case process might well be appropriate, notwithstanding the flaws discussed below. But here, religious exercises were likely to become part of football games in Santa Fe for some time to come. Therefore, rejecting a facial challenge because the policy did not lead inevitably in all cases to prayer and remitting the parties to case-by-case adjudication, as the dissent advocated, would have posed too great a risk to Establishment Clause values.

Case-by-case adjudication could not vindicate the constitutional rights of the student minority. The reason here for distrusting case-by-case review, unlike in the vagueness cases, has nothing to do with the difficulties of detecting unconstitutional discrimination. Rather, the problem was that case-by-case review would do little to actually protect students from religious indoctrination. For that reason, facial challenges were a better means of implementing the Constitution. Leaving the policy in place and waiting to see whether student speakers, in fact, prayed at football games would have “threatened the imposition of coercion upon those students not desiring to participate in a religious exercise.” Courts acting in the case-by-case mode would only review as-applied challenges to the policy after speakers had given their messages, and, in too many cases, after students had been forced to participate in a religious exercise. Whatever relief courts granted would come too late. Facial invalidation thus prevented the coercion of religious minorities

---

189 Id. at 317 (“Simply by establishing this school-related procedure, which entrusts the inherently nongovernmental subject of religion to a majoritarian vote, a constitutional violation has occurred. No further injury is required for the policy to fail a facial challenge.”).

190 Id. at 297 (observing that students had voted for the speaker to give a prayer at the first two football games after the policy took effect); see also id. at 307 (“[A]s used in the past, at Santa Fe High School, an ‘invocation’ has always entailed a focused religious message.”).

191 Thus, the substantiality requirement is implicit only; it is not explicitly enshrined in a constitutional test as in Broadrick, Casey, or Morales.

192 As a general matter, a statute is not subject to a facial challenge because it grants a power that may be abused; instead, we rely on case-by-case adjudication to filter out any abuse of power. A different rule applies here because constitutional violations are likely to be substantial, and case-by-case adjudication is likely to be ineffective. These two considerations make a facial challenge the better means of implementing the Constitution.

193 See Santa Fe, 530 U.S. at 317.
by eliminating, before the fact, the discretion that the policy granted a student majority, rather than waiting for the inevitable unconstitutional use of the power.\textsuperscript{194}

Strategic invalidations also feature in cases in which the Court addresses government aid to parochial schools. In \textit{School District of Grand Rapids v. Ball}, the Court facially invalidated an aid program because religious teachers had too much power to use state funds to inculcate religion.\textsuperscript{195} In the portion of \textit{Ball} that is still good law, the Court invalidated a Michigan school district’s program of community education courses for private schools, virtually all of which were sectarian religious schools.\textsuperscript{196} Under the program, the district sent private religious school teachers, paid with state funds, to teach after-school community education courses at the private schools.\textsuperscript{197} The Court found the program facially invalid because it posed a “substantial risk that, overtly or subtly, the religious message [professional religious schoolteachers] are expected to convey during the regular schoolday will infuse the supposedly secular classes they teach after school.”\textsuperscript{198} Religious teachers simply had too much power and incentive to teach the secular subjects from a religious perspective. Many, if not all, of the professional religious teachers hired to teach the after-school courses “teach in the religious schools precisely because they are adherents of the controlling denomination and want to serve their religious community zealously.”\textsuperscript{199} These teachers “are expected during the regular schoolday to inculcate their students with the tenets and beliefs of their

\textsuperscript{194} Although my claim here is that the excessive discretion strategy often operates differently in vagueness and Establishment Clause cases, that is not an invariable rule. Consider \textit{Larkin v. Grendel’s Den}, in which the Court invalidated a state law granting churches the power to veto a liquor license for an establishment within five hundred feet of a church. 459 U.S. 116 (1982). The Court invalidated the statute because its grant of discretion to churches was an invitation to discriminate.

The churches’ power under the statute is standardless, calling for no reasons, findings or reasoned conclusions. That power may therefore be used by churches to promote goals beyond insulating the church from undesirable neighbors; it could be employed for explicitly religious goals, for example, favoring liquor licenses for members of that congregation or adherents of that faith.

\textit{Id.} at 125. As-applied adjudication was no substitute because the breadth of the discretion conferred on churches made it difficult to monitor the uses of that discretion. \textit{Grendel’s Den} is \textit{Kolender} and \textit{Morales} transposed to the Establishment Clause context.


\textsuperscript{196} \textit{Id.} at 398. \textit{Ball} invalidated two after-school state-sponsored educational programs, which took place almost exclusively at religious schools. Public employees taught in the first program, while teachers from the religious schools taught in the second program. \textit{Agostini} overruled \textit{Ball}’s invalidation of the first program, but did not disturb \textit{Ball}’s holding as to the latter program. \textit{Agostini}, 521 U.S. at 235, 241 n.1 (Souter, J., dissenting).

\textsuperscript{197} \textit{Ball}, 473 U.S. at 373.

\textsuperscript{198} \textit{Id.} at 387.

\textsuperscript{199} \textit{Id.} at 386.
particular religious faiths,” and they teach the state-sponsored after-school classes “in the same religious school classrooms that they employed to advance religious purposes during the ‘official’ schoolday.”

The unfettered discretion religious teachers had to teach from a religious perspective created a substantial risk that religion would, in a great many cases, infuse the state-paid and sponsored after-school courses. Therefore, facial invalidation of the program was necessary, notwithstanding the fact that one of the participating schools was non-sectarian and that the secular courses could, of course, be taught in a secular manner. The Court’s implicit conclusion was that, absent facial invalidation, substantial constitutional violations would likely occur, and these violations would be exceedingly difficult to detect. As-applied adjudication, once again, was insufficient to ensure compliance with the Establishment Clause. The problem, as the Court explained, in discounting the absence of evidence of indoctrination, was that the parties with access to such evidence – the parents, children, and the religious schools – had no incentive to complain about religious indoctrination; in fact, “the students are . . . attending religious schools precisely in order to receive religious instruction.”

As in Santa Fe, facial invalidation prevents religious teaching before the fact by eliminating the discretion-laden program, rather than curbing abuses of such discretion after the fact.

Thus, Ball is another example of the excessive discretion strategy. The Court held the statute invalid on its face not because the law had no permissible applications, but because the risk of unconstitutional state action was too great. Because of the discretion reposed in religious school teachers, the program was an invitation to religious school teachers to use public funds to continue the religious indoctrination that was part of the regular sectarian education provided to students during school hours. It is this grant of...
Key to the Establishment Clause discretion cases is the idea of substantial risk. Facial invalidation is explicitly a tool for risk management. Where, because of the discretionary powers granted, the risk that a statute will be used to advance religion is substantial, and case-by-case adjudication is inadequate to protect religious liberty, facial invalidation protects Establishment Clause values from subversion. The concept of risk, of course, is not unique to Establishment Clause cases, but pervades facial challenges at every level. When facial invalidation is strategic, the practice is a judicial response to the risk of pervasive rights violations. Thus, the risk of chilling protected conduct explains the use of facial invalidation in overbreadth, privacy, travel, voting, and fundamental rights vagueness cases. The risk of arbitrary law enforcement from unfettered police discretion arising from vague criminal laws explains the use of facial invalidation in equal protection vagueness cases. So the Establishment Clause cases are notable not only as another context in which facial invalidation is used strategically, but for their explicit focus on risk management.

Compare Agostini, 521 U.S. at 226 (“[T]here is no reason to presume that, simply because she enters a parochial school classroom, a full-time public employee . . . will depart from her assigned duties and instructions and embark on religious indoctrination.”), and Mitchell v. Helms, 530 U.S. 793, 859 (2000) (O’Connor, J., concurring) (“The Court’s willingness to assume that religious school instructors will inculcate religion has not caused us to presume also that such instructors will be unable to follow secular restrictions on the use of textbooks. I would similarly reject any such presumption regarding the use of instructional materials and equipment.”), with id. at 860 (O’Connor, J., concurring) (stating that “the religious school teacher who works throughout the day to advance the school’s religious mission would also do so . . . during the supplemental classes offered at the end of the day,” and, because the government is paying for those classes, “any religious indoctrination taking place therein would be directly attributable to the government”).

Mitchell comes closest to rejecting Ball, and a plurality of the Court made clear that it would do so. Id. at 826 (plurality opinion). But, as O’Connor’s opinion shows, Mitchell recognizes a basic distinction between an aid program that gives religious school teachers government funds to teach classes without imposing any controls (Ball) and one that merely gives secular educational materials to those teachers, subject to the requirement that the materials be used in a secular manner (Mitchell). In the former instance, the aid statute reposes too much discretion in religious school teachers to advance religion, and this explains why the program in Ball was invalid on its face. In the latter, no substantial risk arises because discretion is twice channeled – once by the fact that the aid is simply a textbook or educational material and, second, by the restrictions on secular uses only.
V. THE STIGMA STRATEGY

The third invalidation strategy – the stigma strategy – calls for facial invalidation in cases in which the law’s effect is to send a message that a particular class of people are second-class citizens. The defect of case-by-case adjudication in such cases is its gradualism. Case-by-case adjudication and its narrowing process take too long; in the meantime, the law’s expressive harms go unremedied. Facial invalidation is, therefore, a better means of implementing the Constitution because it quickly eliminates a law’s stigmatic message.

In Lawrence v. Texas, the Court facially invalidated a Texas sodomy law that criminalized only homosexual sodomy as violative of substantive due process.204 Scholars have been quick to point out Lawrence’s fusion of both liberty and equality values.205 But they have not paid sufficient attention to the precise role that equality norms play in the Court’s opinion. Scholars have not observed the strategic role stigma plays in the Court’s facial invalidation.

In Lawrence, the Court considered both due process and equal protection arguments for striking the Texas law, but the Court chose the broader liberty grounds to redress the expressive harms sodomy laws impose:

If protected conduct is made criminal and the law which does so remains unexamined for its substantive validity, its stigma might remain even if it were not enforceable as drawn for equal protection reasons. When homosexual conduct is made criminal by the law of the State, that declaration in of itself is an invitation to subject homosexual persons to discrimination both in the public and in the private spheres.206

A narrow ruling limited to the particulars of the Texas law – its singling out of same-sex couples – would leave unaddressed the stigma caused by the sodomy prohibition. Moreover, gay and lesbian persons in states with sodomy laws that applied across the board would continue to suffer from the stigma caused by their state’s sodomy laws. If the Court had issued a limited ruling and ducked the continuing validity of Bowers v. Hardwick,207 it might be years before it would have another opportunity. In the interim, gays and lesbians would have to endure the stigma and discrimination wrought by sodomy laws. A ruling on the facial validity of a criminal sodomy law under the Due Process Clause was thus necessary to protect the equality of gay and lesbian persons.208

\[204\] 539 U.S. 558 (2003).


\[206\] Lawrence, 539 U.S. at 575.


\[208\] Lawrence, 539 at 575 (“Equality of treatment and the due process right to demand respect for conduct protected by the substantive guarantee of liberty are linked in important respects, and a decision on the latter point advances both interests.”).
Stigma equally explains why the Court invalidated the Texas law on its face, rather than only as applied to the defendants challenging the statute. After concluding that the Texas statute infringed the liberty protected by the Due Process Clause, the Court suggested a number of constitutionally permissible applications of the law, including instances involving minors, prostitution, and coerced conduct. Notwithstanding the various circumstances in which the Texas sodomy law could be validly applied, the Court held the statute invalid on its face. The stigma strategy explains the Court’s decision. Using case-by-case adjudication to narrow the law to valid limits would have left the law’s stigmatic message in place during the narrowing process, thus exposing gay and lesbian persons to harm in the interim. The Court insisted on facial invalidation because it quickly eliminated the stigma that the law’s declaration of criminality created.

We should take note of the distinct roles that liberty and stigma play in Lawrence. The sodomy statute violates the Constitution because the Fourteenth Amendment protects the liberty to engage in intimate sexual conduct. The Constitution’s protection of the liberty is what renders the sodomy ban invalid in the vast majority of cases. Stigma plays a narrower role in the case and explains the Court’s resort to facial invalidation over as-applied adjudication. Because facial invalidation will quickly eliminate the stigma, it is the better means of enforcing the constitutional right.

The same link between liberty and equality that Lawrence observed may be implicated when other substantive constitutional rights are at stake. Consider English-only laws requiring all government business to be conducted in English, thus forbidding government employees from communicating with persons in languages other than English while performing their official duties. One reason to invalidate such laws on their face is to prevent a chilling effect on protected speech. The English-only law chills too much protected speech between government employees and non-English speakers,

\[209\] Id. at 578.

\[210\] Id. at 579. As in the other examples we have seen, the Court insisted on facial invalidation because the statute was unconstitutional in a substantial number of cases.

\[211\] See id. at 562 (“Liberty presumes an autonomy of self that includes freedom of thought, belief, expression, and certain intimate conduct.”); id. at 578 (“The State cannot demean their existence or control their destiny by making their private sexual conduct a crime. Their right to liberty under the Due Process Clause gives them the full right to engage in their conduct without the intervention of the government.”).

thus necessitating its facial invalidation. Preventing stigma is a second strategic reason for invalidating such speech restrictions. By forbidding public employees to speak in a foreign language, even when approached by an individual who does not speak English, the government sends a message of exclusion – that non-English speakers are not members of the community and that the government is not open to them until they learn to speak English. Even if there are circumstances in which the government may insist that its employees conduct business only in English, facial invalidation is necessary because of the law’s stigmatic harms. Case-by-case adjudication is likely to be too slow, requiring too many as-applied challenges to eliminate the stigma.

In what other contexts does the stigma strategy apply? One might expect the stigma strategy to operate in equal protection facial challenges because of the close connection between stigma and inequality. After all, equal protection doctrine, since its very inception, has been centrally concerned with the prevention of government-imposed stigma. But, in general, the stigma strategy plays no role in equal protection facial challenges. Its principal role is likely to be in fundamental rights cases such as Lawrence, rather than equal protection cases. The reason lies in the tests the Court uses to enforce the mandate of the Equal Protection Clause.

The governing tests used in equal protection doctrine focus on the challenged law’s classification of persons and the reasons the government offers for those classifications. So a statute that classifies based on race or other illicit classification is subject to heightened scrutiny and must be narrowly tailored to a compelling state interest (where strict scrutiny applies) or substantially related to an important government objective (where intermediate scrutiny applies). The doctrinal tests are designed to prevent expressive harms only indirectly, by subjecting laws that make illicit classifications – those most likely to carry stigmatic harms – to a higher burden of justification.

These doctrinal tests lead inexorably to facial review in almost all equal

---

213 See Ruiz, 957 P.2d at 999-1000.
215 In Strauder v. West Virginia, one of the earliest equal protection cases, the Court held that the exclusion of African-Americans from jury service violated Equal Protection because of the stigma imposed; the statutory ban on African-American persons serving as jurors was “practically a brand upon them, affixed by law, an assertion of their inferiority.” 100 U.S. 303, 308 (1880). For discussion of expressive accounts of equal protection, see Matthew D. Adler, Expressive Theories of Law: A Skeptical Overview, 148 U. PA. L. REV. 1363, 1428-38 (2000); Deborah Hellman, The Expressive Dimension of Equal Protection, 85 MINN. L. REV. 1 (2000); Richard H. Pildes, Why Rights Are Not Trumps: Social Meanings, Expressive Harms, and Constitutionalism, 27 J. LEGAL STUD. 725, 754-60 (1998).
protection challenges; as a result, the distinction between facial and as-applied adjudication rarely surfaces.\textsuperscript{218} The reason is quite simple. A statute that contains a discriminatory classification and is not narrowly tailored to serve governmental interests of the highest order will generally be invalid in all circumstances.\textsuperscript{219} In every case, the discriminatory classification triggers and fails heightened or strict scrutiny. For example, in \textit{Gratz v. Bollinger}, every person subject to the Michigan undergraduate admissions policy would raise and succeed on the same as-applied claim; the policy discriminated against him or her on the basis of race and was not narrowly tailored to serve a compelling state interest.\textsuperscript{220} Every time it was applied, the policy violated a plaintiff’s personal equal protection right, making facial invalidation appropriate under \textit{Salerno}.\textsuperscript{221} Case-by-case adjudication, thus, has little, if any, role to play in

\textsuperscript{218} This point holds, as well, for other constitutional rights with a non-discrimination component, such as the Free Exercise Clause and the Dormant Commerce Clause, to name just two. \textit{See infra} note 219.

\textsuperscript{219} \textit{See, e.g.}, \textit{Gratz v. Bollinger}, 539 U.S. 244 (2003) (holding facially invalid under Equal Protection Clause a university’s race-based admissions policy); \textit{Or. Waste Sys., Inc. v. Dep’t. of Envtl. Quality of Or.}, 511 U.S. 93 (1994) (facially invalidating under Commerce Clause statute that discriminated against interstate commerce); \textit{Church of the Lukumi Babalu Aye, Inc. v. City of Hialeah}, 508 U.S. 520 (1993) (invalidating animal sacrifice statute on its face because it impermissibly targeted religious sacrifice, which discriminated against religious practices in violation of Free Exercise Clause); \textit{Kraft Gen. Foods, Inc. v. Iowa Dep’t of Revenue and Fin.}, 505 U.S. 71 (1992) (facially invalidating under Foreign Commerce Clause statute that discriminated against foreign commerce); \textit{Orr v. Orr}, 440 U.S. 268 (1979) (invalidating under Equal Protection Clause statute permitting alimony awards against men, but not women). A similar analysis holds true for statutes that are motivated by a racially discriminatory or other unconstitutional purpose. Such statutes are invalid on their face because every plaintiff can object to the statute because of its unconstitutional purpose. \textit{See, e.g.}, \textit{Hunter v. Underwood}, 471 U.S. 222 (1985) (invalidating criminal disenfranchisement statute motivated by unconstitutional desire to disenfranchise African-Americans); \textit{see also supra} note 178.

\textsuperscript{220} \textit{See Gratz}, 539 U.S. at 251.

\textsuperscript{221} For cases discussing the equal protection right as a personal right, see \textit{Adarand}, 515 U.S. at 224 (“[A]ny person, of whatever race, has the right to demand that any governmental actor . . . justify any racial classification subjecting that person to unequal treatment under the strictest judicial scrutiny.”); \textit{id.} at 227 (requiring strict scrutiny by judges to “ensure that the personal right to equal protection of the laws has not been infringed”); \textit{id.} at 229-30 (“[W]henever the government treats any person unequally because of his or her race, that person has suffered an injury that falls squarely within the . . . Constitution’s guarantee of equal protection.”). \textit{See also} Richard A. Primus, \textit{Equal Protection and Disparate Impact: Round Three}, 117 HARV L. REV. 493, 553 (2003) (explaining that the Court “sees the right to be free from discrimination as every individual’s right to be judged without reference to characteristics – paradigmatically race – that are considered to be morally arbitrary”). For a critique of the personal rights model, see \textit{Pildes, supra} note 215, at 755 (arguing that a “structural conception of rights provides an account, unavailable in the atomistic conception, of why [expressive] harms are constitutionally recognized”). Whether Pildes is correct and the Court should not view equal protection in personal rights terms is beyond the
equal protection analysis. If the statute is invalid to one, it is invalid to all, which explains why equal protection claims are normally adjudicated as facial challenges. Indeed, it is precisely for this reason that facial invalidation is consistent with the Court’s personal rights model, which conceptualizes the equal protection right as a personal, individual right to be free from unjustified racial (or other illicit) classifications.

*Romer v. Evans* illustrates the rare case in which the stigma strategy may be at work. *Romer* is a notoriously difficult case of which to make sense. It may be that *Romer* is yet another example of the general rule that equal protection challenges are adjudicated as facial challenges and that strategic concerns do not enter into the equation. At the same time, much of the discussion in *Romer* suggests the same sort of strategic invocation of stigma we saw in *Lawrence*.

In *Romer*, the Court facially invalidated under rational basis scrutiny Amendment 2, a Colorado constitutional amendment forbidding governmental action to protect gays and lesbians from discrimination. That *Romer* employed rational basis scrutiny – the most deferential scrutiny in equal protection cases – is important. Rational basis review gives great deference to the government’s legislative choices; if the court can find any reasonable connection between the classification and the government’s objective, the court will uphold the statute. Moreover, in rational basis cases, proof of discriminatory motive is unavailing; a statute will be upheld so long as a court can hypothesize any rational basis for the enactment.

*Romer* held Colorado’s Amendment 2 invalid because it swept too broadly;
its “breadth [was] so far removed” from the reasons offered to justify it that the amendment failed rational basis review. The amendment applied across the board and forbade localities from passing any anti-discrimination measure, even in contexts in which Colorado had no legitimate reason to block local action. Colorado’s interest in safeguarding “freedom of association, and in particular the liberties of landlords or employers who have personal or religious objections to homosexuality,” might provide a rational basis for some applications of Amendment 2, but it simply could not justify Amendment 2’s broad sweep. Colorado’s amendment was constitutionally offensive because it preempted every anti-discrimination measure at every level of state and local government. To borrow a First Amendment concept, Amendment 2’s overbreadth was undoubtedly substantial.

But if the problem was the amendment’s excessive breadth – its application to local laws that had nothing to do with the preservation of freedom of association – why didn’t the Court simply reject the facial challenge to the amendment and remit the parties to the process of case-by-case adjudication to narrow the amendment to valid limits, thereby allowing it to take effect in contexts in which it rationally served a valid government interest? Given the deferential nature of rational basis review, case-by-case adjudication seemed to offer a perfectly acceptable solution to cure Amendment 2’s broad sweep.

Moreover, case-by-case adjudication makes some sense if the equal protection right is a personal, individual right as opposed to a structural right. By conceiving of the equal protection guarantee as an individual right, the Court – whether or not it intends to – assimilates the equal protection right to the set of substantive constitutional rights to which case-by-case adjudication has long applied. The personal rights model of equal protection, this suggests, can coexist with case-by-case adjudication. Romer, however, insisted on facial invalidation. The stigma strategy helps explain why.

227 See Romer, 517 U.S. at 635 (describing the government’s proffered justifications for the classification as “impossible to credit” given the breadth of the amendment); see also Roderick M. Hills, Jr., Is Amendment 2 Really a Bill of Attainder? Some Questions About Professor Amar’s Analysis of Romer, 95 Mich. L. Rev. 236, 250-53 (1996) (emphasizing that the breadth of Amendment 2 was at the core of the Court’s invalidation of the measure).

228 See Romer, 517 U.S. at 633 (“[Amendment 2] identifies persons by a single trait and then denies them protection across the board.”).

229 Id. at 635.


231 See Romer, 517 U.S. at 624 (“[Amendment 2] prohibits all legislative, executive or judicial action at any level of state or local government designed to protect . . . homosexual persons or gays and lesbians.”).

232 The obvious alternate account relies on unconstitutional purpose. The purpose account says that Amendment 2 was facially invalid because it was infected with an
In the typical equal protection invalidation, the reason the law is invalid as to any person is the same reason it is invalid as to all. In every case, the law violates the personal rights of those to whom it applies. But in *Romer*, this model does not fit. If we analyze only the case of the employee who fears loss of anti-discrimination protections, we miss the constitutional flaw in Amendment 2. After all, the government might have a valid reason for preempting local employment discrimination measures adding sexual orientation to the list of proscribed bases for adverse action, and thus the amendment might not violate the employee’s personal equal protection right.\(^{233}\)

The reason that Amendment 2 is unconstitutional only emerges on facial review. It is only when the Court considers the breadth of the amendment across cases— not just the employment case, but all cases— that the amendment’s invalidity is apparent. As-applied analysis thus misses the key flaw in Amendment 2 and hides its invalidity.

Thus, Justice Kennedy demands facial review of Amendment 2, refusing to rely on the process of case-by-case adjudication to refine and narrow the breadth with which the Court was so concerned. Justice Kennedy’s move is strategic. Even though Amendment 2 might be rationally related to valid governmental interests in some contexts, Justice Kennedy invalidated the amendment on its face because its very existence stigmatized gays and lesbians as outsiders, as “stranger[s] to [Colorado]’s laws.”\(^{234}\)

Because the amendment, unconstitutional purpose, thereby rendering all applications of the amendment unconstitutional. See Richard F. Duncan, *Wigstock and the Kulturkampf: Supreme Court Storytelling, the Culture War, and Romer v. Evans*, 72 Notre Dame L. Rev. 345, 353 (1997) (surmising that the Court “seems to conclude that the extreme breadth of Amendment 2 suggests an impermissible purpose that permeates all of the Amendment’s applications”); Isserles, *supra* note 25, at 441 n.371 (classifying *Romer* as one of the Court’s “notable decisions [facially] invalidating a statute on the grounds of illegitimate governmental purpose”). But the purpose reading runs into immediate difficulties; namely, the narrow role purpose plays in rational basis cases. In rational basis cases, purpose alone does not suffice to invalidate a statute. See *supra* note 226 and accompanying text. Unconstitutional purpose, thus, cannot explain why, given the over-inclusiveness of the amendment, the proper response is facial invalidation rather than narrowing.

\(^{233}\) *Cf.* Corp. of the Presiding Bishop of the Church of Jesus Christ of Latter-Day Saints v. Amos, 483 U.S. 327 (1987) (upholding validity of provision permitting religious organizations to discriminate in employment based on religion). On the other hand, as-applied challenges might well have been successful in other cases within Amendment 2’s scope. For example, while there might be a legitimate interest in protecting an employer’s freedom to hire whomever he or she wants, no similar justification would support the amendment’s abrogation of police department regulations prohibiting police officers from discriminating against gays and lesbians in enforcing criminal laws.

\(^{234}\) *Romer*, 517 U.S. at 635. Indeed, the Court is quite explicit about the stigma that Amendment 2 causes: “Amendment 2 . . . in making a general announcement that gays and lesbians shall not have any particular protections from the law, inflicts on them immediate, continuing, and real injuries that outrun and belie any legitimate justifications that may be claimed for it.” Id. at 634-35. See also Akhil Reed Amar, *Attainder and Amendment 2*: 
as a whole, caused these expressive harms, the Court refused to remit the matter to case-by-case adjudication to narrow the law, thus invalidating the law even in those applications in which Amendment 2 rationally furthered the government interest in protecting freedom of association. Using case-by-case adjudication to refine the law would have left the amendment’s stigmatization of gays and lesbians in place during a lengthy and gradual narrowing process, a constitutionally intolerable result.\textsuperscript{235}  Facial invalidation was necessary because it would take “extensive adjudications” to narrow the amendment to its valid limits and to eliminate the stigma.\textsuperscript{236} As in \textit{Lawrence}, stigma explains why the Court preferred a facial challenge to case-by-case adjudication.

\textit{Lawrence} and \textit{Romer} share much with the \textit{Kolender/Morales} equal protection vagueness cases.\textsuperscript{237} Both mandate facial invalidation as a strategy to protect persons from discrimination. But the strategy is different in each. In the \textit{Kolender/Morales} line of cases, the Court was concerned that the limitless discretion granted to police officers to define criminal conduct invited discrimination, whereas in \textit{Lawrence} and \textit{Romer}, the stigma the laws created demanded invalidation. The limits of as-applied adjudication are thus different in each line of cases. In \textit{Kolender/Morales}, as-applied adjudication is an insufficient alternative because the risk of discrimination posed by the grant of excessive discretion cannot be checked adequately using as-applied adjudication. The laws’ vague terms make discriminatory enforcement exceedingly difficult to detect. In \textit{Lawrence} and \textit{Romer}, on the other hand, facial invalidation reflects a judicial refusal to tolerate the government’s branding of gay and lesbian persons. Facial invalidation is preferred because it

Romer’s Rightness, 95 MICH. L. REV. 203, 208 (1996) (describing Amendment 2 as a “kind of legal and social outlawry in cowboy country – a targeting of outsiders, a badge of second-class citizenship, a tainting of queers, a scarlet Q”). To borrow Justice Kennedy’s description of sodomy laws, the amendment’s “declaration in and of itself is an invitation to subject homosexual persons to discrimination both in the public and in the private spheres”; its very existence “demeans the lives of homosexual persons.” \textit{See} \textit{Lawrence} v. Texas, 539 U.S. 558, 575 (2003); \textit{id.} at 581 (O’Connor, J., concurring) (“Texas’ sodomy law brands all homosexuals as criminals, thereby making it more difficult for homosexuals to be treated in the same manner as everyone else.”).

\textsuperscript{235} One might object that strategic invalidation seems out of place in rational basis cases, which involve the kinds of discrimination we care least about. It is, of course, true that strategic invalidations in rational basis cases are and should be reserved for exceptional cases. But we need not rule them out entirely. Serious constitutional concerns are present when government sends a message of second-class citizenship to any group, even if the group does not qualify for strict or heightened scrutiny. Such a stigmatizing message is enough to justify, in exceptional cases such as \textit{Romer}, the use of the stigma strategy.

\textsuperscript{236} \textit{Cf.} Baggett v. Bullitt, 377 U.S. 360, 378 (1964) (“It is fictional to believe that anything less than extensive adjudications, under the impact of a variety of factual situations, would bring the [loyalty] oath within the bounds of permissible constitutional certainty.”).

\textsuperscript{237} \textit{See supra} Part IV.A.
eliminates the stigma expeditiously. Even if the statute in Lawrence or the amendment in Romer could be narrowed over time to cases in which they permissibly operated, the process would take too long. In the meantime, the measures would demean and stigmatize gay and lesbian persons and encourage discrimination against them. Facial invalidation prevents the stigma.

CONCLUSION

Beginning with the advent of First Amendment overbreadth doctrine in the 1940s, facial challenges became an important means to protect vulnerable constitutional rights.238 Fearful that sweeping speech restrictions would chill speech and that case-by-case adjudication was powerless to prevent self-censorship, the Court insisted on facial review because it was a better means of implementing the Constitution than as-applied adjudication. Thus was born the strategic facial challenge. Later decades saw the blossoming of similar strategic doctrines throughout constitutional law, as the Court moved beyond overbreadth’s chilling effect theory and developed an array of invalidation strategies.239

Ignoring these developments, Salerno sought to reshape constitutional adjudication by insisting that facial challenges should be the rare exception, not part of the mainstream of constitutional adjudication, and sought to cast doubt on the power of courts to rely on the limits of as-applied adjudication as a basis for employing facial invalidation.240 Salerno created a dual structure that left a space for overbreadth doctrine – but only in First Amendment cases – and rendered the doctrine tenuous and alien to the rest of constitutional law. The problems inherent in that structure became clear almost immediately, as the Supreme Court in case after case (Salerno notwithstanding) used facial invalidation strategically to knock out statutes whose defects could not be cured through case-by-case adjudication without compromising values dear to the Constitution.241 That thread joins Casey, Romer, Morales, Santa Fe, Stenberg, and Lawrence. These cases demonstrate the pervasiveness of strategic facial challenges in constitutional jurisprudence.

Questions of strategy thus loom large in facial challenges. The central question in facial challenges is whether they serve as a better means of implementing the Constitution because of some defect in case-by-case adjudication. When case-by-case adjudication contains no such flaw, a facial challenge should stand or fall based on whether it meets Salerno’s test. But when a case-by-case process threatens constitutional harms, a strategic facial challenge should generally be available. Whether a facial challenge succeeds will generally turn on whether the challenged law violates constitutional rights in a substantial number of cases. Underlying all the forms of strategic

238 See supra Part II.A.
239 See supra Parts III-IV.
240 See supra Part I.
241 See supra Parts III-V.
invalidation is the idea that the rights deprivation must extend to a substantial portion of the class that the law affects. This ensures that the benefits of facial invalidation are high and the costs low.

The key question thus becomes what sorts of cases call for the use of strategic invalidation. The Court’s cases reveal three basic theories: the chilling effect, excessive discretion, and stigma strategies.\footnote{See infra Parts III-V.} By and large, the chilling effect theory pertains to liberty cases, while the discretion and stigma theories deal with different, but closely related, problems of equality. At the root of each theory lies a fundamental commonality: a defect in the process of case-by-case adjudication – some instrumental reason to depart from the use of severability doctrine to separate valid and invalid applications over time. Seen together, these three theories show that strategic protection is an essential part of the law of facial challenges. They provide an important conceptual framework that makes sense of the Court’s facial invalidations and illuminates when courts should invalidate a statute on its face, notwithstanding some valid legal applications.